QUEEN'S UNIVERSITY BELFAST

DOCTOR OF PHILOSOPHY

Human Movement Sonification for Motor Skill Learning

Dyer, John

Award date:
2017

Awarding institution:
Queen's University Belfast

Link to publication

Terms of use
All those accessing thesis content in Queen's University Belfast Research Portal are subject to the following terms and conditions of use

- Copyright is subject to the Copyright, Designs and Patent Act 1988, or as modified by any successor legislation
- Copyright and moral rights for thesis content are retained by the author and/or other copyright owners
- A copy of a thesis may be downloaded for personal non-commercial research/study without the need for permission or charge
- Distribution or reproduction of thesis content in any format is not permitted without the permission of the copyright holder
- When citing this work, full bibliographic details should be supplied, including the author, title, awarding institution and date of thesis

Take down policy
A thesis can be removed from the Research Portal if there has been a breach of copyright, or a similarly robust reason.
If you believe this document breaches copyright, or there is sufficient cause to take down, please contact us, citing details. Email: openaccess@qub.ac.uk

Supplementary materials
Where possible, we endeavour to provide supplementary materials to theses. This may include video, audio and other types of files. We endeavour to capture all content and upload as part of the Pure record for each thesis.
Note, it may not be possible in all instances to convert analogue formats to usable digital formats for some supplementary materials. We exercise best efforts on our behalf and, in such instances, encourage the individual to consult the physical thesis for further information.

Download date: 06. May. 2023

Human movement sonification for motor skill learning

Thesis submitted to the School of Psychology, Queen's University, Belfast, in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

June 2017

John Frederick Dyer, BSc (Hons)

Supervisors: Dr. Matthew Rodger and Dr. Paul Stapleton

Abstract

Transforming human movement into live sound can be used as a method to enhance motor skill learning via the provision of augmented perceptual feedback. A small but growing number of studies hint at the substantial efficacy of this approach, termed 'movement sonification'. However there has been sparse discussion in Psychology about how movement should be mapped onto sound to best facilitate learning.

The current thesis draws on contemporary research conducted in Psychology and theoretical debates in other disciplines more directly concerned with sonic interaction - including Auditory Display and Electronic Music-Making - to propose an embodied account of sonification as feedback.

The empirical portion of the thesis both informs and tests some of the assumptions of this approach with the use of a custom bimanual coordination paradigm. Four motor skill learning studies were conducted with the use of optical motion-capture. Findings support the general assumption that effective mappings aid learning by making task-intrinsic perceptual information more readily available and meaningful, and that the relationship between task demands and sonic information structure (or, between action and perception) should be complementary.

Both the theoretical and empirical treatments of sonification for skill learning in this thesis suggest the value of an approach which addresses learner experience of sonified interaction while grounding discussion in the links between perception and action.

Acknowledgements

My thanks go to my supervisor Matthew Rodger, who has been variously, a guide, teacher, sounding board, therapist and many other things beyond the supervisor's call of duty. His endless patience and support (and willingness to indulge my half-baked ideas) made this thesis possible, and the PhD experience incredibly worthwhile. He deserves far more thanks than I can offer here. Thanks to my second supervisor, Paul Stapleton, for his encouragement and feedback, and for introducing me to a whole world of sonic interaction design and philosophy I would never have seen as a deprived unidisciplinarian.

I would also like to thank everyone who generously donated their time to participate in my experiments - especially the first one! Thanks to my friends and housemates who each did so more than once; for their patience with me during the last few months of this write-up. Thank you to everyone in SARC who helped me out with the practicalities of running several sound experiments when I knew nothing about digital sound. You were vital to this in numerous ways.

I extend my thanks to several people who, while not involved in this research directly, were a positive influence on me and my wellbeing during this time. To my family, particularly my mum and dad, for putting up with a mostly-absent son for the last few months. To so many of my fellow PhD students in the School of Psychology, and in particular the basement-dwellers, who have collectively made studying at QUB a wonderful time that I'll look back on nostalgically for the rest of my life. To a previous research mentor, colleague and friend, Deaglan Page, for many stimulating conversations. To my rock and roll queen, Matthew McCourt, for being a stalwart friend and confidant. You all know yourselves.

To Steph, for being the most supportive partner in the world and always finding ways to cheer me up during the write-up. You are the greatest.

I am grateful for the opportunities I was afforded by the School of Psychology to undertake paid work and engage in activities which greatly furthered my personal and professional development.

I gratefully acknowledge the funding I received from the NI Department for Employment and Learning, without which I would never have been able to take this opportunity.

Thank you.

John Dyer

Queen's University, Belfast

June 2017

ABSTRACT .... **II**

ACKNOWLEDGEMENTS .... **III**

LIST OF FIGURES AND TABLES .... **XI**

CHAPTER 1: INTRODUCTION **1**

CHAPTER 2: PERCEPTION, ACTION AND SKILL **8**

2.1 THE THEORETICAL POSITION OF THE THESIS 8

2.1.1 Benefits of eschewing the traditional approach 9

2.2 PERCEPTION AND ACTION FROM AN ECOLOGICAL PERSPECTIVE 11

2.3 A RICHLY STRUCTURED AND LAWFULLY-RESPONSIVE ENVIRONMENT 14

2.3.1 Ecological information and Affordances 16

2.4 PERCEPTION AS A SKILL 17

2.4.1 Learning to perceive; perceiving to learn 17

2.4.2 Sensory exploration 19

2.4.3 The role of action in perceptual learning 20

2.5 AFFORDANCES AND CONSTRAINTS OF TASKS 22

2.5.1 Constraints on perception 23

2.5.2 Constraints on action 24

2.5.3 Experience in larger-scale tasks and skills 25

2.5.4 Expanding the definition of task and skill 26

2.5.5 Section summary and foreword 28

2.6 FEEDBACK AND MOTOR SKILL LEARNING 29

2.6.1 Intrinsic feedback 30

2.6.2 Extrinsic or augmented feedback 31

2.6.3 Knowledge-of-results 31

2.6.4 Knowledge-of-performance 34

2.6.5 Concurrent augmented feedback 36

2.6.6. Feedback summary 38

2.7 THE ROLE OF AUDITORY INFORMATION IN PERCEPTION AND ACTION 38

2.7.1 *Natural action sounds in rhythmic tasks* ... 39
2.7.2 *Relations between auditory perception and action* ... 40
2.7.3 *Effects of skill in auditory-motor tasks* ... 41
2.7.4 *Altering tasks and task performance with sound* ... 42
2.7.5 *Cross-modal interactions betweenauditory and visual/tactile perception* ... 44
2.8 CONCLUDING REMARKS ... 47

CHAPTER 3: SONIFICATION MAPPING ... 49

3.1 SONIFICATION IN PSYCHOLOGY ... 50
3.1.1 *Mapping movement to sound* ... 51
3.1.2 *The mapping problem in motor skill learning* ... 53
3.1.3 *Error sonification* ... 53
3.1.4 *Statement on error sonification* ... 57
3.1.5 *Direct sonification* ... 58
3.1.6 *Statement on direct sonification* ... 62
3.1.7 *State of the art in Psychology* ... 63
3.2 SONIFICATION IN AUDITORY DISPLAY ... 65
3.2.1 *The problem of mapping complex data to sound* ... 68
3.2.2 *Contextualising the listening task* ... 70
3.2.3 *Auditory 'objects' and 'being' relative to them* ... 73
3.2.4 *Contextualising the meaning-making task* ... 75
3.2.5 *Lessons for motor skill learning from Auditory Display* ... 78
3.3 SONIFICATION IN MUSIC AND INTERACTION DESIGN ... 79
3.3.1 *Intelligence in interaction* ... 80
3.3.2 *Form and function* ... 81
3.3.3 *Embodied interaction styles* ... 82
3.3.4 *Intellectual interaction styles* ... 84
3.3.5 *Mapping movement to music* ... 86
3.3.6 *Sonic constraints on (inter)action* ... 87
3.3.7 *Understanding musical interaction* ... 88
3.3.8 *Interaction summary and lessons for mapping in motor skill learning* ... 93

3.4 SUMMARY: AN APPROACH TO MOTOR SKILL LEARNING WITH SONIFICATION 94

3.4.1 Limitations of the approach 98

3.4.2 Chapter conclusion 99

CHAPTER 4: MOVEMENT SONIFICATION AS FEEDBACK FOR LEARNING A NEW TASK: IMPLICATIONS OF MISGUIDANCE 100

4.1 ABSTRACT 100

4.2 INTRODUCTION 101

4.2.1 Learning and feedback 101

4.2.2 Information presentation and structure 103

4.2.3 Sound and movement 106

4.2.4 The current study 108

4.3 METHOD 109

4.4 RESULTS 119

4.5 DISCUSSION 128

4.5.1 Sonification as a hindrance 129

4.5.2 The guidance effect 131

4.5.3 Unintended task constraints 131

4.6 EXTENDED DISCUSSION 134

4.6.1 Task conceptualisation and mapping decisions 134

4.6.2 Learning bimanual coordination tasks 138

4.6.3 Augmented feedback without transformation 141

4.6.4 Concluding thoughts on participant experience 143

CHAPTER 5: TRANSPOSING MUSICAL SKILL: SONIFICATION OF MOVEMENT AS CONCURRENT AUGMENTED FEEDBACK ENHANCES LEARNING IN A BIMANUAL TASK. 144

5.1 ABSTRACT 144

5.2 INTRODUCTION 146

5.2.1 Movement sonification and the guidance hypothesis in motor skill learning 146

5.2.2 A perception-action perspective on the guidance effect in bimanual tasks 148

5.2.3 *Noisy events, perceptual unification and sonification* *148*

5.3 METHODS 152

5.4 RESULTS 158

5.5 DISCUSSION 160

5.5.1 *Benefits of sonification for performance in acquisition* *161*

5.5.2 *The 'guidance effect' in early retention* *163*

5.5.3 *Exploiting the musicality of movement in retention* *164*

5.5.4 *Conclusion* *166*

CHAPTER 6: ADVANTAGES OF MELODIC OVER RHYTHMIC MOVEMENT SONIFICATION IN BIMANUAL MOTOR SKILL LEARNING AND LISTENING FOR EXTENDED RETENTION 167

6.1 ABSTRACT 167

6.2 INTRODUCTION 169

6.2.1 *Information pickup in motor skill learning* *169*

6.2.2 *Potential benefits of a richer melodic mapping* *172*

6.2.3 *Prolonging motor skill retention through musical listening* *173*

6.3 METHODS 176

6.5 RESULTS 181

6.6 DISCUSSION 183

6.6.1 *Performance benefits with the use of melody* *184*

6.6.2 *Learning and refresh with musical learning* *186*

6.6.3 *Motivation* *188*

6.6.3 *Caveats and conclusions* *189*

CHAPTER 7: SONIC INFORMATION DESIGN CONSIDERATIONS: INFORMATION OR KNOWLEDGE AS AUDITORY FEEDBACK FOR MOTOR SKILL LEARNING 191

7.1 ABSTRACT 191

7.2 INTRODUCTION 192

7.2.1 *Sonification and motor skill learning* *192*

7.2.2 Mapping movement to sound ... *193*

7.2.3 Problems with sonifying the measure variable ... *194*

7.2.4 Guidance effect implications ... *195*

7.2.5 Burdensome intellectual elaboration ... *196*

7.2.6 The current experiment ... *198*

7.3 METHOD ... 201

7.4 RESULTS ... 205

7.5 DISCUSSION ... 209

7.5.1 Using the sonified metric ... *209*

7.5.2 Metrics and music in retention ... *211*

7.5.3 Limitations of the current study ... *213*

7.5.4 Conclusion ... *214*

CHAPTER 8: GENERAL DISCUSSION ... 216

8.1 EMPIRICAL REVIEW ... 216

8.2 IMPLICATIONS AND RECOMMENDATIONS FOR SONIFICATION MAPPING DESIGN IN MOTOR SKILL LEARNING ... 220

8.2.1 The perils of metaphorical sonification ... *220*

8.2.2 Sonification for musical skill ... *223*

8.2.3 Knowledge, information and variable selection in sonification ... *225*

8.2.4 Improper variable selection ... *226*

8.2.6 Section summary ... *229*

8.3 LIMITATIONS AND POSSIBLE EXTENSIONS OF THE RESEARCH ... 229

8.3.1 Scope ... *230*

8.3.2 Situated skill in the use of sonification ... *230*

8.3.3 Moving with sound, but not sonification ... *232*

8.3.4 Transfer effects ... *233*

8.3.5 Applications outside the lab ... *234*

8.4 CONCLUSION ... 235

BIBLIOGRAPHY ... 237

APPENDIX A: VIDEO RECORDINGS OF SONIFIED TASK PERFORMANCE 266

List of figures and tables

Figure 3.1: A taxonomy of interactions, as perceived by spectators. 91

Table 3.2: Interaction styles with a highly constrained instrument. 93

Figure 4.1: Experimental apparatus as seen by the participant. 111

Figure 4.2: Experimental apparatus as seen from above, with markers highlighted. 111

Figure 4.3: Task schematic for shape manipulation task. 113

Figure 4.4: Global temporal error over the course of practice and retention testing for three experimental conditions. 120

Figure 4.5: Regression R^2 values (representing smoothness of cylinder movement) for practice and retention. 122

Figure 4.6: Mean time difference (s) between cylinder arrival and the arrival of the contemporaneous right-hand shape. 124

Figure 4.7: Mean positional error of arrivals/stops on the first movement of each trial for three feedback conditions across practice and retention testing. 127

Figure 4.8: Lissajous figure example, from Ronsse et al., (2011) 140

Figure 5.1: Bimanual shape-tracing apparatus 153

Figure 5.2: Terminal feedback graph 154

Figure 5.3:.Melody produced by correct performance 155

Figure 5.4: Experimental procedure for experiment 2. 157

Figure 5.5: Rates of average absolute ratio error for the three feedback groups during practice, retention and transfer in experiment 2. 158

Figure 6.1: The demonstration melody in experiment 3. 177

Figure 6.2: Rates of average absolute bimanual ratio error per condition in pre-test, practice, retention testing and transfer testing for each of the three training condition groups for experiment 3 181

Figure 7.1: Schematic of the experimental procedure in experiment 4 204

Figure 7.2: Rates of average bimanual ratio error for each experimental condition over pre-test, practice trials, retention testing and transfer in experiment 4. 206

Chapter 1

Introduction

This thesis is about skill, sound and human technology. More specifically, it is about how people can use computer-generated sound as a supportive aid for learning new skills. ‘Naturally-occurring’ sounds, which we hear every day, are relatively easy to understand as they can tell us something about what is going on around us. The sounds of footsteps, voices, or something crashing to the ground nearby are meaningful features of everyday life. Computer-generated sounds, however, can sometimes be strange, otherworldly and difficult to interpret (a quality historically well-exploited by the science fiction genre of film and TV). This (potential) disconnect between sound and ‘real-world’ events is at the heart of the attempts made in the current thesis to understand the use of computer-sound for learning new skills.

Sonification – turning into sound – is a conceptual bridge between the world and artificial sound (Hermann, Hunt, & Neuhoff, 2011). The modern version of this practice takes digital data gathered through measurement of some quality of the world and turns it into sound, so that something which might not have been audible before can be listened to. The aim is that sonification can help a listener to make new discoveries or achieve new understanding through hearing the data. A recent high-profile example of this can be found in the research output of the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the USA. The physicists, upon their discovery of gravitational waves (so-called ‘ripples in spacetime’), immediately sonified their data waveform. The result, an audible ‘chirp’, which *sounds* like the feeling of a bump, or ripple, is equivalent to listening to the sound of two black holes merging, millions of light-years away[1]. Listening to the sonified data, it is easy to get a sense of the dynamic structure of the event one is observing, despite the esoteric nature of how the data were gathered and the underlying theory of general relativity. Listening to data bypasses the need to look at graphs and read about lasers and calculations, as well as opening up the experience of discovery for those with visual impairment.

The notion of exploring the world through sonification can also be applied in settings closer to home. Metal detectors, for example, allow beachcombers to locate hidden metal objects despite their invisibility beneath the sand. Parking sensors on the rear of many modern vehicles (in which the rate of heard beeps signals proximity to an object) give drivers an extra source of sensory information which can aid in the difficult task of parallel-parking. With practice, the use of these devices can feel like sensory augmentation, as if we have been endowed with new powers of perception (Clark, 2003).

The performance of a range of human skills is dependent on sound perception, to a greater or lesser extent. Music and dance can involve some of the most spatio-temporally precise and complex motor coordination seen in any domain of human activity. Musicians and dancers move their bodies with respect to perceived structure in sound, and in the case of music-making, create sound which is directly informative about the underlying movements. Physical practice in these domains would not make sense at all to do without sound – it is vital to the activity itself. Success and error are primarily *heard*. Very fine-grained differences between the timing of events can be perceived by a listening human, beyond what is possible using the other senses. This can enable movements to be coordinated on an equally fine-grained time-scale – with practice. However, even very well-practiced motor skills – like walking or running, for example – which might not seem to be dependent on hearing one's movements, can be disrupted if their associated sounds are manipulated (Tajadura-jiménez et al., 2015). A wider discussion of how sound and movement interact in motor tasks will be undertaken in a later chapter.

---

[1] Listen to the sound of gravitational waves here: http://podcasts.nytimes.com/podcasts/2016/02/11/science/space/ligo-chirp/LIGOChirp.mp3

There has been some interest in recent years in using sonification to make human movement audible where it might otherwise either be silent, or produce vibrations below the level of human hearing. The aim is that computer-generated sound can be useful to people learning new skills, in the same way as physically-produced sounds are useful to musicians or runners. In this case, the data sonified would be data about movement, or a person's interaction with the environment, and sound would be fed back 'live' - during movement - rather than later, as in the dataset sonification described earlier. Sonification of human movement is already sometimes applied in a sporting context – enabling a new, more temporally acute means of performance-monitoring. With enough practice using the system, mistakes in movement performance can be heard – and perhaps heard more easily than they could be seen. Elite sports teams are beginning to take notice of sonification as a potential tool for performance enhancement. Several examples of sonification applied in a sporting context have already been published. In swimming for example, the pressure of water against the hands in breast-stroke has been sonified, so that swimmers can hear and fine-tune the propulsive force they exert (Schmitz et al., 2013). In rowing, athletes have been trained with sonification of oar trajectory, boat acceleration and pulling force (each in a different implementation), with the aim that their performance might be tuned towards greater biomechanical efficiency (Effenberg, Fehse, Schmitz, Krueger, & Mechling, 2016; Schaffert & Mattes, 2015; Sigrist et al., 2011).

There is also emerging medical interest in sonification as sensory augmentation in motor rehabilitation (Rosati, Rodà, Avanzini, & Masiero, 2013). Individuals who have experienced certain kinds of stroke can lose motor function on one side of the body. Sometimes perception of movement on the same side is degraded as well. Sonification of arm movement has seen some success in small-scale trials, in which the patient practices simple movements (like reaching for an object) with their impaired arm. The sound produced by movement can stand in for missing or degraded sensory feedback which would otherwise be informative about the reach (Schmitz, Kroeger, & Effenberg, 2014; Scholz et al., 2014). In Parkinson's disease (PD), sufferers find it difficult to coordinate movements in rhythmic tasks, like walking. It has been established for several years that music, or even a rhythm can cue walking in PD patients. Recently however, sonification has been successfully implemented (Rodger & Craig, 2016; Schiavio & Altenmüller, 2015). These medical applications are still at the early stages, but show great potential for movement recovery.

While sonification can clearly work as a performance aid in some motor tasks, it is not necessarily guaranteed that it will work well in every new case. There are some published examples of movement sonification which do not enhance motor performance, despite being designed to do so. The relationship between movement and sound produced by the sonification system – termed the 'mapping' – is identified in the current thesis as a critical factor in the effectiveness of sonification. Detailed consideration of the mapping raises endless questions. For example, in a reaching task (typically a skill to be recovered in stroke rehabilitation), should the position of the hand be tracked and sonified? Perhaps the angle of the elbow as it opens? Beyond that, what kind of sound should the data be mapped to? Would the pitch of a tone, increasing with elbow-angle work? What information does the patient need the most? Will they recognise the most important auditory cues when they occur? Does the patient need to *like* the sound, or will it suffice to simply make the information available? Will the skill stay learned when the sound is turned off at the end of the session? Questions like these and many more arise upon consideration of any new task to be sonified. This is where the earlier-mentioned disconnect between real-world events (movement in this case) and computerised sound becomes both an opportunity and a problem. Digital sound synthesis techniques enable the production of infinite variety of sound, which can be controlled by any movement data as input. This provides considerable freedom to the designer of the sonification as to how to map movement to sound. However, due to the relative novelty of movement sonification as a technique, there is little guidance on how to create the best mapping, unless the task has been sonified and successfully tested before.

The mapping between movement and sound has not received much theoretical attention in the research literature up to this point, and comparisons between different kinds of sonification are rarely reported. Where they are reported, the types of sonifications compared are not always clearly motivated by theoretical considerations, i.e. it can be difficult to know why one version of sonification led to better performance on the motor task rather than another. Lack of reporting on theoretical guidance for mapping decisions can turn a successful sonification report into a standalone result, impossible to replicate in other tasks.

The aim of the current thesis, broadly, is to add to our understanding of how and why sonification of movement works to enhance the learning of motor skills. A clear model of what successful movement sonification *does* can help to constrain mapping design choices. For this research project, the following, more specific research questions were formulated:

1) How do listeners understand the information provided by sonification?
2) How should designers characterise the relationship between movement and sound (the mapping)?
3) Does providing extra movement feedback through sound lead to dependence *on* sound, or can sonification be used for learning, and then safely removed?
4) To what extent does the kind of sound used - and how it is mapped to movement - matter?

To address questions 1 and 2, the current thesis examines the links between perception and action in motor skill learning, then considers how artificial sound should be designed to work within these processes. Research in this area of Psychology has tended to reduce questions of perception, feedback and learning to their neural substrates and left vaguely-defined concepts like listener experience and motivation in the background. It is my contention that these are important factors which have a causal role in motor performance enhancement with sonification. This thesis aims to incorporate literature from other disciplines in the Arts and Humanities which have addressed these critical issues more directly. The theoretical approach to mapping developed in this work both informed and is informed by the empirical investigations reported in chapters 4-7, which together address questions 3 and 4 above. The theoretical approach of this thesis can be taken to exist within an 'embodied' approach to the mind, which takes the agent-in-the-world as its unit of analysis (Varela, Thompson, & Rosch, 1991). The empirical work utilises high-frequency optical motion capture to collect data for movement performance analysis and to be used in digital sound synthesis.

The thesis is organised into two theoretical chapters, which set the background for four subsequent empirical chapters, then a final chapter of general discussion.

In chapter two, a theory of motor skill learning is established in which perception and action are conceptualised as part of a whole system which incorporates the agent and their environmental niche. This chapter also reviews previous work on the use of augmented feedback in motor skill learning. Finally, the effects of sound on motor performance are considered, including how sound can be an assistive aid for movement.

In chapter three, three discrete disciplines which use sonification are visited. In Psychology, I discuss the various ways in which artificial sound has been mapped to movement to enhance the learning of new motor skills and highlight a lack of consistent methodology and guiding theory. The field of Auditory Display is concerned primarily with the kind of sonification exemplified in the LIGO’s gravitational wave sonification - that of datasets into sounds. I use this section to explore how listeners can extract knowledge about the world from sound when the world is not there to help. In the final section, I describe current research in computer-music interaction philosophy, which helps construct a broader picture of sonic interaction as a socio-culturally embodied system.

Chapter four reports the first empirical study of the thesis, in which a tabletop object-manipulation task is sonified to enhance learning of a sequence. Participants learn the task with 1) sonification of all parts of their movements, 2) only the end-points of each movement, or 3) no sound. In an extended discussion, I highlight the crucial importance of information structure in motor skill learning.

The fifth chapter reports a sonification experiment with a different motor task (complex bimanual coordination) in which the dependency of learned skills on sonification is assessed. Additionally, the effect of sonifying only a demonstration of good performance is tested, to verify that there is actually an advantage of using live, movement-controlled sound.

In chapter six, the question of information structure is addressed in an experiment. Here, participants learn the same bimanual task as chapter 5 with two different kinds of sonification. This experiment aims to clarify the relationship between sound morphology and the movements involved in the task. The limits of motor skill retention are tested and a procedure devised to extend retention after performance has declined.

Chapter seven reports the final experiment of the thesis, in which the question of information structure in sonification is addressed from a different angle. Participants learn the bimanual coordination task with one of two kinds of sonification, each designed with a different conceptualisation of information/knowledge, and different ideas of the way in which sonification can be useful to a learner.

Chapter eight discusses the empirical results in the context of the perception-action approach to skill and interaction established in chapters two and three. Some recommendations for how sonification designers should map sound to movement are provided. Limitations in the proposed theoretical framework for sonification are highlighted, including ways in which the empirical chapters left aspects of the theory in need of further research. Suggestions for how this research program could continue and some targets of application for the research are discussed.

Chapter 2

Perception, Action and Skill

This chapter elaborates a theoretical approach which takes the agent, environment and the relations set up and maintained between the two as a foundation for the analysis of skilled motor performance. It begins by making the case for taking a perception-action approach in more detail, then examining the underlying theory and assumptions more closely. Following that, a complementary theoretical framework for perception as an active skill is fleshed out. The second half of the chapter deals with historical and current research in the study of motor learning and feedback, before exploring the particular influence of auditory information on the perception-action system.

2.1 The theoretical position of the thesis

In this thesis, discussion of sound, motor skill learning and use of interactive systems is conducted within what can be loosely characterised as a 'perception-action' approach. While traditional approaches to these topics in Psychology have called upon internal mechanisms and representations for explanation, the perception-action approach entails a commitment to explaining behaviour at the behavioural level. Inspired by James Gibson's Ecological approach to Psychology, a fundamental tenet of this approach is the mutually-supporting relationship between perception and action. On the one hand, perception is for action and can be thought of as purposeful and task-oriented. One the other, perception is an active process which requires movement to generate information. The perception-action approach is eminently suitable for dealing with questions of interaction around sonification as feedback for motor skill learning, in which sound informs the learner about their ongoing action. The approach set out in this chapter seeks to explain behaviour and phenomenological experience in terms of interactions between brain, body and environment. Therefore, it has strong links with the emerging embodied approach to the study of the mind (e.g. Varela et al., 1991).

2.1.1 Benefits of eschewing the traditional approach

Literature dealing with motor skill learning in Psychology generally exists under the large theoretical umbrella of what might be called the ‘information-processing’ or ‘cognitive’ approach. This approach holds that human experience of the world is 'indirect', and perception is *of* an internal representation coded in neural activity, constructed from noisy, impoverished sensory data and given form by the ‘top-down’ application of knowledge (Fodor & Pylyshyn, 1981). Thought and learning is defined computationally, as the rule-based manipulation of mental constructs, like a computer program might manipulate stored variables (Fodor, 1975). Information-processing approaches to motor skill learning and motor control follow this pattern by building conceptual models which describe what the motor system might be doing if it were assumed to be a computer. Despite often providing a good fit for behavioural data, these models are purely hypothetical constructs, initially created to stand in for processes which were themselves not directly observable (Cisek & Kalaska, 2010; Verwey, Shea, & Wright, 2015). This began to change with the advent of brain-imaging technologies, as neural activity and plastic reorganisation became observable during learning (e.g. Asaad, Rainer, & Miller, 1998; Imamizu et al., 2000). As neural correlates of learning were found, the effort to merge computational models with neuroscience began (Albright, Kandel, & Posner, 2000). An issue with tying computational models directly to brain activity however, is that it forces highly abstract models, devised without consideration to the specific real-world machinery which might enact them, to become mechanistic models, which exist *in* brain tissue (Weiskopf, 2011). This effectively confines to the brain what might otherwise be conceptualised as a wider process, distributed across brain, body and environment. As part of a wide-ranging critique of traditional cognitive theory in Psychology, Hendriks-Jansen (1996) reminds the reader that computational models should be seen as models of behaviour, not models of brain or mind.

Cisek and Kalaska (2010) identify a range of difficulties encountered in the merging of cognition and neuroscience where motor skill learning and motor control are concerned. Researchers sometimes find that computational models do not fit with empirical observations of activity in neural tissue. Central of these is the finding that brain areas involved in the (as traditionally conceptualised) separate processes of perception, cognition and movement show substantial overlap in activation (Culham & Kanwisher, 2001; Lebedev & Wise, 2002).

The fundamental notion of representations in the brain has been critiqued at length in recent years (Chemero, 2009; Kiverstein & Miller, 2015; Lakoff, 2012), paving the way for an 'embodied' version of Psychology which gives greater consideration to the ways in which the body and its interaction with the environment can structure cognition. Taking a wider view of cognition as a process which draws on a distributed set of resources allows for inclusive and meaningful discussion of things which are important, possibly constitutive parts of the process but are themselves not located solely in the brain. Music and musical aesthetics, for example, may not find the most elegant or useful explanation in purely neuroscientific terms. We could arguably best understand the place of music in human Psychology by taking an Anthropological perspective, which might consider music's inherent connectedness with bodily movement, a setting, social collaboration and perhaps its evolutionary fitness implications (Mithen, 2005). At this grain of analysis it may be possible to better characterise the activity of music-making itself, possibly thereby assisting with the interpretation of associated neural data (Altenmüller, 2007). Furthermore, with due attention granted to brain-body-environment dynamics and a description of the broader task in which the agent is engaged, models of behaviour could be built which stand up to scientific and logical scrutiny, while also providing some approximation of phenomenological experience (Kaufer & Chemero, 2015). As this thesis deals with issues surrounding the use of interactive sound systems, including information structure, experience, skill, design, aesthetics and meaning - an embodied account is, I believe, worth pursuing. The theoretical underpinnings of this thesis will be explored in more detail in this and later chapters. It is expected that the theoretical framework adopted and developed here will help further understanding of sonification and how it might best be deployed in service of motor skill learning.

2.2 Perception and action from an Ecological perspective

The Gestalt Psychologists of the early 20th Century argued that there were already-structured sources of sensory information in the environment related to objects and events which could be perceived holistically (Humphrey, 1924). In other words, for the Gestalt Psychologists, the basic units of perception were whole objects and dynamic events, rather than unstructured points of sense data which needed to be combined and enriched through internal processing for perception to take place. This dovetailed with contemporary work by phenomenologists such as Martin Heidegger (and later, Maurice Merleau-Ponty), who recognised that the experience of perception is immediate, and that breaking it down into its constituent stimuli is usually an abstraction away from experience. The idea of direct perception was expanded by James Gibson (1950). Gibson's major development was to show how Gestalt visual-perceptual grouping principles were related to the movement of the perceiver over time. Following his study of how WWII pilots learn to land planes, he conceived,

> "*...the possibility that there is literally no such thing as a perception of space without the perception of a continuous background surface.*" (Gibson, 1950, p. 6).

This simple insight has far-reaching implications for the study of perception generally (and by extension, action). In other branches of Psychology, research into visual perception aims to explain how the brain constructs our sense of three-dimensional space from two-dimensional optic information hitting the retina. This sense-making Cognitive view is predicated on the separation of perception and action into distinct Psychological processes (modules), overseen by a central integration process. In opposition to this separation, Gibson links all three processes inextricably together. In the above example, Gibson describes the phenomenological experience of space as an emergent property of the relationships between objects in the visual field, and how those relationships vary lawfully over time as the perceiver moves.

Within this theoretical perspective, perception and action are conceptualised as a coherent, whole system which incorporates several interacting parts – human movement itself, a lawfully-responsive, structured environment, and the two in concert actively creating the information necessary to coordinate further successful behaviour. Motor behaviour is often conceptualised as a loop, spanning brain, body and environment (Witt, Linkenauger, & Wickens, 2016). Most of the 'computation' in the previous example is embodied in the dynamic expression of lawful relations between the parts of the system. The job of the agent is to become a skilful perceiver, sensitive to higher-order relationships in the optic array (like, the lawful co-variation of background and foreground objects based on distance and degree of movement) and use them to coordinate successful behaviour (Pick, 1992). Computationally, it is less demanding for a brain *not* to have to represent a concept of space, when it (or, the agent) could instead learn the relationships and thereby make ‘space’ directly perceivable (Warren, 1998).

The above example, in which the pilot must move relative to the environment to perceive space, underscores one of the fundamental tenets of Gibson’s approach to Psychology: perception-action mutuality. This entails, firstly, that perception is purposeful. It is *for* the coordination of action. Secondly, that action is necessary to generate the information required for perception. Gibson and his followers wrote primarily about visual perception, but the general perception-action approach to Psychology they espouse can be applied across modalities, including audition (J. J. Gibson, 1966; Steenson & Rodger, 2015; Turvey, Shaw, Reed, & Mace, 1981).

A complementary theory of motor control is proposed by Bernstein (1967). Bernstein noted that the human motor system technically contains more room for output variability than is necessary for most motor tasks. Each moving body part contains several possible dimensions of movement, or degrees of freedom (DOF), for deployment in the performance of a motor task. The fact that there are more possibilities for movement than is necessary to complete a task effectively has often been characterised as ‘redundancy’, or ‘abundance’ (Latash, Scholz, & Schöner, 2002). In a kinematic analysis of a hammering factory worker, Bernstein describes how the various joints and segments of the worker's body move in relation to each other during work. On every hammer-strike, there is enormous variability in the DOF, but almost no variability at all in the meeting of hammer and metal. The challenge for a moving individual, as Bernstein saw it, is to identify the combinations of muscular patterns (among these abundant DOF) that will consistently and stably achieve task performance goals. In other words, the problem is to map the enormous range of limb movements that *can* be made to the narrower range of movements that will lead to successful performance of the task at hand. Bernstein's intuition was that the factory worker does not control the exact deployment of all the DOF, despite their functional role in the achievement of task goals. Wulf and Shea (2002) make the case that a focus on the lower levels of motor control (i.e. exact deployment of the DOF) is attentionally demanding, and is therefore a detriment for both complex skill acquisition and subsequent performance. The argument is that the individual could never possibly control - or ever learn to control - every part of the system. Todorov and Jordan (2002) propose a model of motor coordination to account for this, in which motor output is constrained only according to the goals of the task, and control is exerted over higher-level sensorimotor contingencies. Variability in redundant dimensions is not only a by-product of this arrangement, but precisely what allows for adaptive behaviour -that is, behaviour which can be stable under varying conditions and resistant to perturbation. In fact, they argue that skilful performers exploit DOF variability in order to take advantage of the opportunities which might open up *during* movement. Examples of this behaviour can be found in expert cello-bowing (Verrel, Pologe, Manselle, Lindenberger, & Woollacott, 2013), skiing (Vereijken, Emmerik, Whiting, & Newell, 1992) and writing (Newell & van Emmerik, 1989). This proposed solution to Bernstein's DOF problem pairs well with the Ecological approach to perception, in that it provides a model which explains why the individual not only does not *need* to plan movements ahead of time, but also how they benefit from not doing so.

2.3 A richly structured and lawfully-responsive environment

One of the requirements for Gibson's theory of direct perception is an environment which responds in a consistent manner (Michaels & Carello, 1981). In other words, information at the point of perception must be specific to the object or event which precipitated it. Concerning visual information, light reflects off an object in a way which specifies its physical qualities. This relationship between objects and the way in which they structure the optical array is described as an 'invariant', i.e. something which does not vary under changing points of observation. Michaels and Carello identify two kinds of invariants. The first, structural invariants, are the ways in which physical features of an object itself can lawfully structure the optical array. For example, objects of the same size and shape (say, cubes made of chalk and coal) structure the patterning of light reflected off them in different ways.

> "*This means, simply, that the correspondence between the structured light and the surface composition, size, shape, position, and other characteristics of the object or place are derivable from the laws of physics*." (Michaels & Carello, 1981, p. 23)

Under the same lighting conditions, a cube made of chalk always reflects more light than one made of coal. Michaels and Carello acknowledge that the pickup of structural invariants might be difficult to quantify (especially when more complex object features are considered, e.g. texture), but it is possible to understand in principle how this specifically-structured information might be useful to a perceiver without necessarily doing so.

A second kind of invariant is relevant to understanding how perception occurs in a dynamic coupling with the environment. An encounter with an object in the environment is an event which happens over time, and for information to specify an event, the event itself must structure energy arrays over time in a consistent manner.

> "*Besides the changes in stimuli from place to place and from time to time, it can also be shown that certain higher-order variables of stimulus energy – ratios and proportions, for example – do not change. They remain invariant with movements of the observer and with changes in the intensity of stimulation… They constitute, therefore, information about permanent environment. The active observer gets invariant perceptions despite varying sensations*." (J. J. Gibson, 1966, p. 3)

Michaels and Carello (1981) term these "transformational invariants". Pickup of this kind of invariant information is what allows an agent to perceive consistency in the world around him/her and coordinate movement with respect to landmarks. Gibson himself identified several such invariants, including for example, 'motion parallax', the invariant relationship between distance from perceiver and rate of change of visual angle with lateral movement – the same invariant which enables pilots to perceive distance and land safely on the ground. Lee (1976) identified a similar higher-order informational variable to account for how drivers visually estimate the time-to-arrival of an object, in the form of the ratio of optical angle and optical expansion rate over time. By picking up this variable (dubbed *tau*) directly and controlling its rate of change over time, it is possible to detect and modulate the time-to-contact between oneself and an object in the visual field. Gaver (1993) argues that the acoustic array is structured specifically according to the dynamics of events in the same way as the optical array. Auditory perception can therefore also be described in terms of structural and transformational invariants. As examples of structural invariants, the sounds made by noisy objects in the environment can allow a listener to perceive things about its quality, for example, its material composition (Giordano & McAdams, 2006), or its size (Sedda, Monaco, Bottini, & Goodale, 2011; van Dinther & Patterson, 2006). Auditory transformational invariants can be identified by characterising how events structure the acoustic array over time, e.g. a moving object can specify its position and speed of movement, allowing it to be intercepted (Houben, Kohlrausch, & Hermes, 2004), or, in the case of a moving vehicle, avoided (Gaver, 1993).

Perception research within the Ecological tradition aims to discover the often complex, high-level informational invariants which matter for an individual who is performing a task.

2.3.1 Ecological information and Affordances

Following Gibson's line of argument, what is perceived is not sense data, but objects and events themselves - which are already meaningful items of perception in the primary instance. While structured patterns of stimulation in physical media (i.e. what has been termed 'information') enable direct perception, they are not in themselves sufficient. van Dijk, Withagen and Bongers (2015) argue that the term 'information' as used by Gibson does not imply that it carries intrinsic content, or meaning in itself. Perception, in the sense advocated by the Ecological approach, only actually happens upon co-specification of an agent with the complementary 'effectivities' (skills, experience, goals etc.) to make use of information for the coordination of action. This idea will be developed further in the following section.

According to Gibson, we perceive 'affordances' - the opportunities for action which the environment provides (J. J. Gibson, 1972). Objects and events are therefore perceivable in terms of the activities they afford. For example, a glass affords grasping by an agent with an appropriately shaped grasping apparatus (an arm and hand) and the ability to coordinate a reach-to-grasp action. The exact definition of affordances has been (and continues to be) the subject of debate in the Ecological Psychology literature, due to the fact that Gibson did not initially define the concept as clearly as the field has come to require (Chemero, 2003; Rietveld & Kiverstein, 2014; Stoffregen, 2003). This debate revolves around the question of whether affordances are objective, perceiver-independent properties of the environment or relational properties of the perceiver-environment system. Most theoretical camps agree that affordances can only be perceived/actualised with the specification of an agent/animal with the corresponding effectivity to perform the action solicited by the given property of the environment. To use a classic example, a step may only afford stepping up if a perceiving agent is of the appropriate height (Warren, 1984). Whether the step can be said to have the property of 'step-ability' outside of a situation in which that agent is present with the intention of climbing, is still under discussion.

2.4 Perception as a skill

This section will explore mechanisms of perception at the behavioural level. To a large extent inspired by the writings of the Social Anthropologist, Tim Ingold, this line of argument will attempt to place discussion of both 'perception' and 'action' under the terminological umbrella of 'skill' (Ingold, 2000, 2001). Ingold contends that the study of skill necessitates taking an ecological approach, which helps to advance a view of skill in its material and cultural context (Ingold, 1996, 2001, p. 21). The thesis of this section is that adaptive task performance (alternatively, skill) is behaviour which is enacted in a richly-structured environment via learned, task-appropriate strategies of perceiving and acting.

2.4.1 Learning to perceive; perceiving to learn

If information is abundant and already richly-structured according to its source (Michaels & Carello, 1981), then a major component of learning must be refinement of perception. Eleanor Gibson (1969, 1988), who can be credited with developing much of the theory of perceptual learning, called this process ‘differentiation’. There is an enormous amount of subtle (and not-so-subtle) variation in the information which impinges on the sensory receptors at all times. Early in perceptual development, much of this variation is missed. Perceptual learning is conceptualised as the development of the ability to differentiate one kind of sensory stimulation from another. J. J. Gibson and Gibson (1955) illustrate perceptual learning by differentiation with the example of two wine tasters. There are many thousands of types of wines in the world, all with different chemical signatures. One novice taster might be able to discriminate between red wine, white wine, sherry and champagne in a blind taste test. “He has four percepts in response to the total possible range of stimulation.” (p. 35). However, a connoisseur of wine might be able to discriminate between several kinds of red wine, even differentiating between types of grape, region of origin and vineyard, all by taste. This taster might have thousands of percepts in response to the “total possible range of stimulation”, i.e. all possible wines he or she might imbibe. In line with the ecological approach to perception being concurrently developed by James Gibson, it was argued that this difference in perception was a learned ability to differentiate ever more subtle variations in the rich information arriving at the sensory organs, in this case, how different wines impinge differentially upon receptors in the tongue, nose, mouth and throat. The information for detecting fine differences in wine is present in an encounter with the wine; the taster has only to learn to detect it.

Much of the early experimental work which underpins the notion of differentiation made use of just such a perceptual discrimination paradigm, in which participants were required to discriminate between a set of near-identical of stimuli. J. J. Gibson & E. J. Gibson (1955) tested participants’ ability to recognise a particular swirling scribble on a card from a deck of 32 cards bearing similar but subtly different scribbles. Some were very noticeably different to the target scribble, for example they could be mirrored, or the scribble could be shrunk substantially. However, some differed only very slightly from the target. They could, for example, appear slightly more tightly/widely swirled. The Gibsons found that simply attempting to identify the target among the cards in the deck several times, without feedback about correctness, led to improvement in performance on the identification task. This simple finding was particularly novel (and has become a classic in the decades since) because participant performance was improved without the use of any extrinsic feedback strategies. At the time, it was widely believed that improvement in perceptual ability could only take place by building associations, or with the top-down application of knowledge about the world and (in this case) knowledge-of-results feedback about judgement correctness (Adolph & Kretch, 2015). In contrast, this experiment showed that simple experience and motivated attention could induce changes in information pickup ability, leading to more accurate and fine-grained perception. It implies that a wine connoisseur can be trained if he/she is motivated to explore the sensory array and try to pick out the differences between wines (for an example of successful perceptual learning in beer tasting, see Lelièvre-Desmas, Chollet, Abdi, & Valentin, 2015).

2.4.2 Sensory exploration

The importance of purposeful exploration for perceptual learning cannot be overstated, as perceptual learning does not happen without agent action (E. J. Gibson, 1988). A tenet of the Ecological approach to Psychology is that perception and action are mutually interdependent processes which serve each other. In other words, while perception may be *for* action, action is also required to differentiate information for perception. In this way, perception can be understood as a bodily skill, a way of knowing how to successfully interact with the world (Noe, 2004).

Gibson originally thought that agents learned to differentiate information for its own sake, as demonstrated by her work on static visual displays (E. J. Gibson, 1969). This experimental paradigm used mostly abstract stimuli, and might have given rise to the interpretation that perceptual learning happens through discrimination of features of the environment per se, which are potentially differentiable primarily due to the nature of how physical features structure ambient energy arrays (Michaels & Carello, 1981). Later, Gibson became more convinced that perceptual learning was related to the ways in which differentiated information supported variety of action, i.e. that agents learn to perceive affordances (E. J. Gibson, 1988; Pick, 1992).

There is potentially infinite undifferentiated information impinging on the various sensory receptors at all times, most of which remains undifferentiated. To differentiate information and have meaningfully distinct perceptions requires information be used for something. For example, one might not necessarily learn to differentiate between very similar hues of paint unless motivated to do so, either by the demands of an experiment (Ozgen & Davies, 2002), or by virtue of being engaged in a bedroom-redecoration project.

Such a scenario would entail particular attention paid to the differences between hues, for the purpose of making a judgement on which hue might be more appropriate as a wall covering – or to provide an answer to the experimenter. Outside these scenarios, this information would not be differentiated; ‘eggshell’, ‘ivory’ and ‘lily’ paint, despite having different optical properties, would collectively be ‘white paint’. Similar differentiation occurs in sound perception. Phillips-Silver and Trainor (2007) presented participants with an ambiguous auditory rhythm with no accents, which could be interpreted as having either a double or triple-beat. They found that practicing a certain kind of bodily movement along with the un-accented stimulus rhythm (bouncing by bending the knees) in time with either the second or third beat biased later similarity judgements in favour of the rhythm moved to. Participants who had bounced on every third beat rated the sound of the initial stimulus rhythm as more similar to a rhythm with an accent every third beat - the opposite pattern was observed in participants who had bounced on every second beat. Similarly, Su and Pöppel (2012) demonstrated that rhythmic entrainment is facilitated by prior movement of the listener's body with an ambiguous rhythm (using whatever style of movement was preferred, e.g. head-nodding, foot-tapping). What this demonstrates is that structured sensory stimulation can be perceived differently depending on what the perceiver can use it for. Gibson would eventually argue that information is differentiated by being brought into use for action. Or, put another way, one perception is different from another inasmuch as it affords a different action.

2.4.3 The role of action in perceptual learning

Action has a constitutive role in perceptual learning. Eleanor Gibson and colleagues observed that infants, when encouraged to approach their caregiver across a variable surface, tended to inspect and palpate the surface with their hands before travelling across it – or not travelling across (Adolph, Gibson & Eppler, 1990, as cited in Pick, 1992). In some cases, the surface was found to be solid, and in others, a deformable and undulating water-bed. This task, locomoting to a caregiver across unfamiliar terrain, requires action on the part of the infant to make information available which specifies the nature of the surface, and whether or not it might afford crossing. The infants who had already learned to walk were observed to explore the situation for much longer than habitual crawlers, and to adjust their locomotion style to the conditions. That is, walkers opted to crawl across a waterbed surface rather than walk, whereas crawlers simply crawled after only brief haptic exploration. The sensory consequences of exploratory action here enable perception of the property of 'supportability' - whether the surface will support walking across. This experiment further demonstrates that the particular actions employed to bring action-relevant information into use are dependent on the specific needs and 'effectivities' of the infant at their stage of development and level of locomotive skill. In this example, haptically-perceived properties of the surface are much more relevant to the infant who can walk than the infant who can only crawl. Therefore, the information specifying a waterbed versus a solid surface is differentiated, and brought into use to support correspondingly separate behaviours. Gibson contended that the emergence of task-appropriate, skilful behaviour was the product of perception and action capabilities building upon each other incrementally. Infants learn to perceive their environment by actively exploring and interacting with it, thereby learning via experience the actions which reliably produce information for perception, and the contexts in which certain strategies are effective. In a discussion of how children learn to carry objects, Gibson said,

> "*Carrying is especially interesting to the developmental psychologist who wishes to relate detection of new affordances to developing cognition because it suggests a spiralling process, beginning with perception of the simplest affordances, such as separability and contactability, then moving on to chewability and graspability, then to reachability, to hideability, and eventually to all the refinements of transportability. With each new coil of the spiral, new properties of surfaces, objects, and events are perceived as consequences of exploratory activity, building an ever richer cognitive world. Detecting new affordances provides the means of differentiating the properties of things.*" (E. J. Gibson, 1988, p. 34)

Gibson contended that complex skills, like carrying, develop out of simpler elements, like being able to perceive that an object is solid enough to be grasped. Competent performance of simpler skills (including differentiation of the information that is made available as a consequence of exploration) supports even wider exploratory behaviour with the simpler skill serving as a scaffold for the detection of more information, and the perception of higher-level affordances. As the infant discovers that touch generates information for the perception of surface properties, so too do toddlers discover that a bed can be bounced upon to generate further visual information and perhaps enable perception of objects outside of a high window, each with affordances of their own - and so on. In this way, some affordances are only perceptually available to the highly skilled. For example, a cello might afford playing the music of Bach for a highly skilled cellist, but perhaps only plucking the strings to a complete novice or young child (Moneta & Csikszentmihalyi, 1996; Robinson, 2011). James Gibson and his followers see task specific skills, or ‘effectivities’ and ‘affordances’ as co-specifying – i.e. an affordance can only be perceived as a possibility or ‘actualised’ if the agent is appropriately skilled in the perception-action task in question (J. J. Gibson, 1972; Turvey et al., 1981).

2.5 Affordances and constraints of tasks

> "*[i]nformation is specific to the ecological **task** of the animal. It is not specific to mechanisms or to processes within the animal, nor is it purely external unrelated to the organism*" (Reed, 1996, p. 57, emphasis mine).

The above quotation suggests that in a given task, there should be information available to an agent which supports performance. There may also be some informational variables which are maximally useful, and it might be that 'experts' are those who can detect and use them. Learning to perform a task and becoming skilful can therefore perhaps be characterised as a process of 'attunement' to the most-important features of the task, which entails a constant mutual development of perceptual and movement abilities (Newell, McDonald, & Kugler, 1991).

2.5.1 Constraints on perception

Information is selected from the rich array sampled by the sensory organs primarily to serve performance of the actions which are appropriate given the situation or task in which the agent finds themselves. In this way, not everything in the environment is perceived simultaneously; indeed, Simons and Levin's famous demonstrations of 'change-blindness' (i.e. the phenomenon whereby an individual engaged in a primary task often does not perceive changes in their visual scene which appear, from the observer's privileged perspective, to be meaningful) show that people tend to perceive quite selectively, and in a way which is constrained by the situation (Simons & Levin, 1997). Rensink, O’Regan and Clark (1997) argue that the 'gist' of a visual scene (or, alternatively, the 'gist' of a task), can be determined quickly if the task in question is already familiar to the perceiver, and some of the objects featured have, for him/her, some use-relevant meaning. Using photographs of everyday scenes in a 'spot-the-difference' task, Rensink et al. show that changes to (participant-rated) 'important' objects are recognised much more quickly than changes to 'unimportant' features. A change made to a relatively unimportant feature of the visual scene, for example, the positional shift of a railing behind a couple eating outside a cafe, required, on average, many more alternations between the original and altered image, occurring over a longer period of time, before it was identified by participants. On the other hand, a change to an 'important' feature, such as the position of a helicopter flying low to the ground, required only very few alternations, over a very short time.

Rensink et al. (1997) suggest that the orientation of attention to meaningful features of the task underlies the phenomenon of change blindness. Their participants were already, in some sense, experts - at perceiving everyday scenes. As such they have habitual ways of perceiving and acting. A lifetime of experience interacting with everyday situations means that some features of an everyday visual scene immediately show up as more meaningful than others - the couple rather than the railing, for example. For the purposes of the task at hand, which the framing of the static photograph implies is *'observing the couple'* and not '*walking straight through the scene'*, the railing shows up as little more than background, and is not directly attended to.

2.5.2 Constraints on action

Similarly, the movements performed in order to generate information are likely to be constrained by both the task itself and level of task-relevant skill. In other words, experts at a task should (obviously) move differently, and in doing so, bring different information into use than should novices. Reingold, Charness, Pomplun and Stampe (2001) provide an elegant demonstration of this task-action-specific expert advantage in a change-blindness study, using chess players of varying skill and employing chess-related stimuli. They tested memory for piece positions at different skill levels, and also saccadic and foveal fixation behaviour. As expected, experts showed different patterns of eye movement compared to novices for pickup of chess-related positional information. They made fewer individual fixations per trial than novices, and used a strategy of fixating between sets of pieces, rather than upon single pieces. It appeared that experts were perceiving relations between pieces rather than the position of each piece - a more efficient strategy which made use of higher-order, contextually-specific informational variables. This indicated that attaining chess expertise entails developing a strategy of perceiving the "necessary interpiece relations from both foveal and parafoveal regions" (p. 54). At first blush this might sound as if chess experts, in becoming experts, have come into possession of a generalised ability to pick up information from a wider visual span than novices. However, this was a task-specific skill which was present only when pieces were meaningfully configured according to the game - and disappeared when pieces were scrambled.

This interpretation of the change-blindness effect is intended to illustrate the idea that perception is not general and context-free; neither is it a passive, reactive or automatic process that happens within us upon exposure to a stimulus. Rather, it is an intentional activity performed in and constrained by the task or situation. What is perceived is what is attended to – and knowing what to attend to is a learnable bodily skill[2] (Noe, 2004). To become an expert in a task, an agent needs to become adept at differentiating information related to features to which he/she would need to be maximally responsive. In chess, this could be the relations between pieces. On a first date, the facial expression of one's partner. In a game of rugby, the emergence of 'passable' gaps in the opposing line (Correia, Araújo, Cummins, & Craig, 2012). Making some use of such information is itself what makes perceptions meaningful. Making effective and adaptive use of information is the hallmark of skill.

2.5.3 Experience in larger-scale tasks and skills

There exists a growing body of research into differences between what is perceived by experts and novices in tasks requiring a greater degree of bodily movement. It would be expected, based on the position advocated so far, that informational variables attended to and used in service of task completion would change as expertise is gained. Fowler and Turvey (1978) characterise motor skill learning as a movement towards, in perception and action terms, attunement with task-critical informational variables. This incorporates orienting towards, differentiating and selecting relevant information for the online control of action. In baseball, Castaneda and Gray (2007) show that expert batting performance is facilitated by drawing attention to higher-level features of the batting task as a whole (such as the ball leaving the bat), rather than the physical execution of the swing (hands and arm movement) - which conversely, benefitted novices. In a wide-ranging review, Wulf (2013) provides many similar examples of manipulations to attention in sports and motor skills which follow the same pattern for experts vs. novices. These findings indicate that task-specific skill is a

[2] It is easy to forget that information pickup from static experimental stimuli is as action-dependent as in an interactive, more ecologically-relevant task. In order to perceive some features of the scene and not others, individuals must move intentionally so as to generate information for perception. This might entail bodily orientation, optical focussing, a series of saccades and fixation on the objects which one has learned to be responsive to. This reflects, as Ingold (2001) put it, a “fluency and dexterity of […] movement” (p. 27).

product of attending to the information which best supports performance, taking into account the agent's current ability to make use of it.

Upon this basis, we can reasonably speculate that the phenomenological experience of playing chess is different for individuals at different levels of proficiency (see: Dreyfus, 1996 for just such an analysis). An encounter with a chess board is demonstrably characterised by entirely different perceptions and actions at expert level compared to novice (Reingold et al., 2001). This is an approach which can be applied to all manner of skills; experts move differently, and thus, perceive differently (and vice versa). Ingold (2001) applied such a perception-action analysis of skill to bag-making among the Telefol people of New Guinea and also nest-building in the male weaverbird to illustrate the continuity of this approach. Expert bag-makers also had a particular way of moving, which was contingent on a perceptual engagement with emergent visuo-tactile information in the context of the task. Ingold supposes that this type of task analysis is a way to formalise something like what is colloquially referred to as ‘the feel’ of a skill. In my view, one of the major strengths of an ecological analysis of skill is that it provides a compelling mechanism for subjective experience. If we know the skills and experience of an agent, we might also thereby know (or at least speculate about) what features of a situation stand out as important to him/her, and how he/she might be solicited to act and what information he/she might thereby bring into use. In the words of Thomas Nagel (1974), it lets us talk sensibly about “what it is like to be…” an experiencing agent performing a task, or in a situation.

2.5.4 Expanding the definition of task and skill

Until now, this chapter has mainly dealt with tasks of the kind which are amenable to laboratory study, in which the actions and perceptions of the agent are mostly measurable (and will largely continue to do so). However, some authors in the Ecological Psychology community propose that the same conceptualisation of skill developed here can be ‘scaled up’ to accommodate a much wider range of activities, typically within the purview of other branches of Psychology concerned with ‘higher’ cognition (Bruineberg & Rietveld, 2014; Ingold, 1996).

> “*There is no need at all to limit engagement with affordances to a limited set of motor skills (e.g., grasping a cup, climbing stairs, sitting on chairs). The variety of affordances available to us as humans is as rich and varied as the abilities and sociocultural practices we are socialized into as human beings through processes of ‘enskillment’ which take place in already structured material surroundings*.” (Rietveld & Kiverstein, 2014, p. 343)

These authors build upon Ingold, who suggested that there are learned ways of behaving which might be thought of as culturally-based – or in which the ‘correct’ way to behave is only specified by interpersonal/societal convention. However, these activities too can be deemed ‘skilful’ in continuity with more concretely-situated perception-action tasks (Ingold, 1996). Where exactly one might draw a boundary between ‘lower’ perception-action skills and behaviour in adherence to ‘higher’ cultural conventions in the first place is already very unclear. For example, blind and partially-sighted individuals coordinate their locomotion in urban environments with respect to the vehicle sounds available in the street (Koutsoklenis & Papadopolous, 2011). This is clearly a learned perception-action task enacted for the purpose of safe navigation in urban environments when one is unable to see where the road is. However, this skilful behaviour only makes sense in the wider cultural context of urban life and the system of rules which govern the behaviour of drivers.

Rietveld and Kiverstein (2014) suggest that certain affordances are only available in a “form of life” (p. 327). James Gibson pointed out that all animals alter their environment to be more supportive for their actions (e.g. digging burrows, building nests, making trails in long grass), and that human beings are no different (J. J. Gibson, 1986). In the above urban example, we have altered the environment a very great deal (constructed a city and instituted a civil and legal apparatus to ensure that we all drive within certain pathways), and thus a form of life has emerged. For this reason, roads afford driving for the qualified driver, and pavements do not. Pavements afford walking for the pedestrian, and roads do not. Neither pays much attention to the goings-on in the other's world until they must briefly merge (e.g. a pedestrian wishes to cross the road, or a driver loses control of his/her car). Both driver and pedestrian perceive and act in a way which is constrained by the conditions of their immediate task – included in these conditions is the tacit acknowledgement that each lives in an industrialised society where there are certain rules for behaviour.

A culture of practice can also be seen as a constraint on perception and action, and as a real, constituent feature of a motor task. Take musical practice, for example. For a Czech listener, the sound of an accordion, tuba and guitar might afford dancing a polka. To a Swiss listener, similar physical stimuli might afford yodelling. As another example, there is nothing about the physical structure of a drum which constrains the rhythm with which it can be struck, except the musical landscape in which a would-be-drummer has been enculturated.

An individual agent may be selectively sensitive to some features of a task, or predisposed to interact with a task in a certain way because he/she has learned that some ways of behaving are conventionally practiced in such a situation.

2.5.5 Section summary and foreword

What has (hopefully) crystallised out of the current section is an acknowledgement of the importance of sensory and motor engagement with the supportive structures of the environment for the development of skill:

> "*Development involves changes in animals' bodies, perceptual sensitivity, action capabilities, and environments*." (Adolph & Kretch, 2015, p. 129)

The aim of the current section has been to develop a conceptualisation of skill which can be carried forward and used to guide further discussions of perception and action in the context of motor tasks and sonification. In summary, the environment offers potentially infinite opportunities for action. In the process of becoming skilful, a learner may explore these options and learn to differentiate some information and movement strategies as more useful than others for the achievement of task goals. Over the course of development, the learner may discover more useful, or higher order informational variables which support further improvement; action strategies may also change with reference to new information. Skill is specific to features of the task, including task goals, supportive structures available and sometimes, the broader socio-cultural context.

Given that it is likely to be a very time-consuming process, might it possible to speed up the development of skill through some intervention? The following section will address a mostly separate area of experimental literature, related to the use of artificial feedback to enhance the learning of skills. Historically, empirical investigations of skill acquisition with feedback have not taken the approach outlined so far; indeed, even the prevailing use of the terms 'motor learning' and 'motor skill learning' in the literature speak to the tendency for skill to be seen as primarily an issue of motor execution - of simply moving the right way. Performance is typically measured at the behavioural level, by tracking kinematics or by defining an outcome variable, and 'learning' is characterised as the improvement of the ability to bring those variables within an acceptable range. An extrapolation of the theoretical perspective taken so far in this chapter suggests that artificially-generated information has enormous potential - to guide action where key task-intrinsic information may be lacking or difficult to perceive. Where possible, an interpretation of experimental findings commensurate with this perception-action view will be provided.

2.6 Feedback and Motor Skill Learning

Feedback, in the most general sense, is information made available through action in a responsive environment. When feedback is perceivable, it is the primary means by which agents judge the success (or otherwise) of their actions and identify errors in motor performance. Thus, feedback is fundamental to the process of motor skill learning. The experimental literature on feedback in motor skill learning recognises several major categories of feedback, which will be explicated in this section. In some cases, the boundaries between categories of feedback are not clear-cut; there can be category overlap in a given implementation, and traditional descriptors do not neatly fit in all cases.

2.6.1 Intrinsic feedback

A basic division in the literature on feedback which is particularly relevant to the current thesis is that between 'intrinsic' and 'extrinsic' (sometimes termed 'augmented') feedback (Magill, 2011). Intrinsic feedback, as the name suggests, is information which is available to a moving individual as an intrinsic part of the task in which he/she is engaged. The environment responds in predictable ways given certain motor activity. This link between human action and its environmental outcome allows such information to be distinguished from background information unrelated to motor performance (Holst & Mittelstaedt, 1980). Intrinsic feedback is often classified in terms of the sensory modality in which it is picked up (Ernst & Bülthoff, 2004). For example, picking up a glass of whiskey might induce perceivable changes in visual, auditory, olfactory, proprioceptive and haptic energy arrays, which would be picked up by corresponding sensory organs sensitive to those energy types (eyes, ears, nose, muscle spindles and mechanoreceptors respectively). 'Intrinsic' feedback information arises out of the interaction between agent and environment, its form constrained by familiar laws of cause-and-effect (Michaels & Carello, 1981). Perception of performance using this information is achieved without any extrinsic human/technological mediation (although practice is may be required). To make more sense of this last point, it is helpful to remember that intrinsic feedback is defined in opposition to its counterpart, extrinsic feedback (this will be the subject of the following subsection).

Using intrinsic feedback to judge the quality of movement performance entails perception of movement outcome relative to a goal state. Detecting differences between the form of movement produced and the target form can enable identification of errors in performance, leading to their correction and gradual refinement of performance with practice (Ericsson, Krampe, & Tesch-Römer, 1993). However, motor skill learning (generally defined as relatively permanent performance improvement on a task) with intrinsic feedback alone can be difficult and progress may be slow, as 1) the most useful task-intrinsic information may not be immediately obvious, 2) intrinsic feedback might not necessarily make errors any more salient than non-errors (or vice versa), and 3) intrinsic feedback does not in itself specify what ideal performance should look like.

2.6.2 Extrinsic or augmented feedback

Feedback meets the definition of 'extrinsic' if it is generated by an additional system to - and does not arise naturally from - the immediate task (Schmidt, 1991). This is feedback which is provided via some sort of external mediator or display[3]. Extrinsic feedback is frequently described as 'augmented' feedback because it is intended to be *more useful* than intrinsic information sources alone. Unlike intrinsic feedback, extrinsic feedback might not necessarily conform to the familiar laws of cause-and-effect one might expect from non-technological systems and the ‘natural world’. In most cases, this is either a result of human intercession (e.g. verbal commentary from a coach), or transformation carried out by man-made machinery (typically digitisation and computation), which can elaborate almost any conceivable kind of information based on measurement of performance as input (Sloboda, 1986).

2.6.3 Knowledge-of-results

Given the lack of restriction on form imposed by the use of human and human-made systems to deliver feedback, there is enormous variety in augmented feedback solutions. However, several categories can be identified in the literature, through which the usefulness of augmented feedback can be systematically evaluated.

Some of the earliest research on feedback in motor skill learning was concerned with the effect of simple verbal information about outcomes on task learning. Following the example of Thorndike, (as cited in J. A. Adams, 1971) participants learning a new motor

[3] Although the term 'display' implies a screen, this is not necessarily the case. Information can be 'displayed' in different formats, including visually, though sound, physical/haptic machinery, or some combination.

skill might be given rather generic qualitative verbal feedback such as "good" or "bad" following a practice attempt. Alternatively, feedback might be quantitative, and more specific to the task. For example in golf, a coach might say, "Your shot went into the right rough", or a sprinter might look at their time on a scoreboard (examples taken from Magill, 2011, p. 334). These forms of feedback aim to communicate knowledge of the result of a practice attempt. From this information, a learner should be able to tell whether what they did on the previous trial was successful, and sometimes, the degree of success achieved. Knowledge-of-results feedback (widely termed KR) was applied in a lever displacement task by Bilodeau, Bilodeau and Schumsky (1959). Participants were required to learn to move a lever to a predefined position, while receiving KR about the degree and direction of positional error. In this task, the goal position was not visible or demonstrated to participants, forcing reliance on KR as feedback. Results showed that participant performance improved when they were given access to KR information. The results of this pioneering research (along with similar studies from the same era) provided some of the first empirical demonstrations of motor skill learning using primarily extrinsic, declarative knowledge about outcome, rather than intrinsic sensory information (for a review of early research in this area, see J. A. Adams, 1971). In general, it seemed that the more KR was provided, the greater the benefit to motor skill performance (Bilodeau & Bilodeau, 1958; Newell, 1974).

In later studies, participants were not forced to rely solely on KR as a primary source of task-relevant information. Several studies tested the benefits of KR when participants were introduced to the task by a demonstration, which was repeated regularly throughout practice (Lai & Shea, 1998; Schmidt, 1991; Winstein & Schmidt, 1991). Additionally, researchers began designing experiments which included retention testing without KR feedback, following a sustained critique of existing methods of conceptualising performance versus learning by Salmoni et al. (1984). In one typical example, Winstein and Schmidt (1991) asked participants to practice and learn a lever manipulation motion with a specific pattern of amplitude over time. In one experimental condition, feedback was provided at the conclusion of every practice trial. This consisted of a visual trace of the participant's movement amplitude over time, overlaid on a graph of perfect performance. Additionally, the root mean square error (a value representing deviation from correct performance) for the trial just completed was displayed simultaneously with the trace. Participants who received KR on every practice trial were able to use the extra information to acquire the motor skill more quickly in the early stages of practice than those in a lower KR-frequency condition (33% or 50% of trials with feedback), however performance differences had disappeared by the end of practice. In no-feedback retention, results showed that - contrary to prior research showing that more assistance is better than less (Adams, 1971) - performance was substantially better in the low KR-frequency conditions in retention.

This finding was replicated several times (Schmidt, 1991; Schmidt, Young, Swinnen, & Shapiro, 1989), leading to the formation of a hypothesis that, in general, there was an inverse relationship between the percentage of practice trials enhanced with extrinsic feedback and performance on a no-feedback retention test. This became known as the 'guidance hypothesis' or 'guidance effect', based on the notion that participants must be relying too heavily on the guidance provided by augmented feedback. This hypothesis also acknowledges the importance of intrinsic sources of sensory information for learning (in particular, proprioception, given that much early research studied the effect of only intrinsic proprioception + KR on performance). The benefit of limiting KR was hypothesised to work similarly to the 'specificity of learning' hypothesis, which states that the effect of training is greatest when conditions between acquisition and testing are closely matched, and that recall is to some extent tied to the conditions of encoding (Barnett, Ross, Schmidt, & Todd, 1973). Withholding KR on some trials forces participants to attend to intrinsic feedback during practice, which means that conditions between practice and no-feedback retention testing are less dissimilar.

The inclusion of a retention test without feedback and identification of the guidance effect reflects a general refinement in both theoretical scope and methodology in this area in the late 20th Century. From early reviews of the effect of KR on learning (J.A. Adams, 1971, 1987), it is clear that motor skill learning had not yet been clearly distinguished as a discrete subfield within the broad theoretical umbrella of learning and memory. J.A. Adams (1971) proposes that KR in these early experiments advanced a view of ‘skill learning’ as a memory-based problem-solving task, in which the participant must discover the solution (correct motor output) from indirect clues given by the experimenter. This definition allowed for a very wide range of activities to be investigated using similar problem-solving methodology, including some which today would be considered purely intellectual in nature, e.g. language skills, mathematical techniques and a doctor making a clinical diagnosis (examples from J. A. Adams, 1987, p. 42). The aim of this research paradigm, as articulated by Seligman (1970), was close to that of the behaviourist tradition – a search for general principles of learning which could apply across all domains of human activity. For some time, Psychologists did not draw a distinction in kind between the theorised knowledge structures for semantic learning and those for motor skill learning. In time, experimental methodologies evolved to reflect a narrower interest in *motor* skill learning, which came to be seen as dependent on underlying mechanisms particular to the domain of controlled movement, rather than general mechanisms of learning. This refinement in scope had the effect of orienting research questions toward those which were more applicable to performance outside the lab.

2.6.4 Knowledge-of-performance

In non-laboratory motor skill learning situations, coaches do much more than inform their student about error rate, or deliver scores. From the learner’s perspective, relying only on knowledge-of-results is an extremely inefficient method to progressively limit potentially infinite variability in motor output (i.e. to improve performance). KR prescribes no adjustment to the quality or kinematics of movement performance itself, only an adjustment to the outcome. Bernstein (1967) identified the effectively infinite possible interactions between connected body parts (each with their own axes of movement) as a problem for a task novice. However, KR does not guide the learner in constraining the choice. With KR, the criterion outcome might be somewhat abstracted from motor performance itself, for example, the score from the judges in pool diving, or the lap time for a sprinter. Learning with KR in the lab is often a case of trial and error, in which the learner has to do a large amount of intellectual work.

Following a trend towards investigating something closer to ‘real-life’ motor skill learning and aiming for greater ecological validity in lab-based motor skill research generally (Abernethy, Thomas, & Thomas, 1993; Salmoni et al., 1984), researchers have also been interested in the potential benefits of more direct feedback about movement quality. ‘Knowledge-of-performance’ (or, KP) feedback is information about how the skill or movement was executed, in other words, that which *led* to the result. To revisit two earlier examples, in golf, an example of KP might be, “You did not take your backswing back far enough before you began your downswing.” In sprinting, the sprinter might watch a video replay of a race they previously ran (examples from Magill, 2011, p. 334). This is information which can be used to communicate specific corrections to motor performance.

It has been found in general that KP+KR is more effective for acquisition of a new skill than KR alone. In basketball shooting, Wallace and Hagler (1979) found that participants who were given verbal feedback about their stance and bodily kinematics (plus KR) made more successful shots in practice and no-feedback retention testing than participants who only received KR. Kernodle and Carlton (1992) argue that KP is especially effective for learning more biomechanically complex motor skills (those requiring the coordination of many degrees of freedom) after finding benefits of KP over KR in throwing. Although these examples use a human coach as a delivery mechanism, KP might not necessarily be verbal, or even verbalisable in the same way as KR. For example, a visual graph of limb displacement over time would count as KP, given that it displays information about how the movement itself was performed (Vander Linden, Cauraugh, & Greene, 1993). However, sometimes the distinction between KR and KP is not easy to make. In a paper mentioned previously (Winstein & Schmidt, 1991), participants had access to a visual trace of their pattern of movement over time, which, on its own, would be KP. However this trace was overlaid on a graphical form which would be produced by perfect task performance – the goal state. Participants were therefore shown whether or not they had produced the goal of movement, and how close to the goal they had come. Because "the movement goal was isomorphic with the movement pattern" (Winstein & Schmidt, 1991, p. 679), the authors conceptualised this visual display as KR, although other authors have done the opposite with similar displays (e.g. Vander Linden et al., 1993)[4].

2.6.5 Concurrent augmented feedback

Here, it is useful to highlight yet another set of subcategories of augmented feedback. Augmented feedback may be delivered terminally (post-trial, after movement has finished – as it has been in nearly all studies reviewed in this section so far), or concurrently (during and alongside motor performance). The nature of KR usually precludes concurrent delivery (a result/outcome is required), but the same is not necessarily true for KP. KP can be delivered while a practice attempt is in progress. For example in archery, a coach might say, "Your elbow is too low," during which time the learner is still lining up the shot and has opportunity to adjust his/her performance. Alternatively, a graphical display showing a kinematic pattern can be updated live to show how a practice attempt is progressing, allowing adjustments to be made 'online' and on the basis of what it shows (Vander Linden et al., 1993). The distinction between KP and KR continues to blur in the case of concurrent augmented feedback. In a wide-ranging review, Sigrist, Rauter, Riener and Wolf (2013) identify many implementations of concurrent feedback delivered in the visual, auditory and haptic modalities – many of which undoubtedly provide knowledge of performance via the live display of kinematic variables – but which also provide a kinematic target expressed via the same display.

[4] In this kind of lab-friendly task, achievement of the task goal can be expressed in terms of movement kinematics, as the task itself is to move in a certain way. This reflects a latent conceptualisation of motor skill as production of motor output. Many real-world tasks having a primary outcome variable which is abstracted from kinematics (e.g. drive distance in golf, making a shot in basketball, etc.).

To resolve this ambiguity, it might be necessary to recognise something subtle yet distinctive about how concurrent augmented feedback has come to be implemented which sets it apart from both KR and KP as traditionally deployed. Today, augmented feedback is usually displayed concurrently with movement. This might mean, for example, that when a wrist is flexed, a display plots hand displacement by the movement of a point which leaves a trace behind (Kovacs & Shea, 2011). The pattern displayed is under the direct control of the learner, in a moment-to-moment fashion. The learner is then able to watch the display, as it is updated live, with a representation of his/her performance. This is similar to everyday movement (i.e. outside the lab), in which information specifying the state of the (task/motor) system is immediately responsive to the agent's actions. Given that concurrent feedback displays are coupled to action in a rule-based, predictable way (such that exact reproduction of the previous trial's movement will give rise to physically identical sensory feedback), relationships between perception and action can be learned, and concurrent feedback can be used to coordinate action in a way not dissimilar from intrinsic feedback – i.e. by making use of perceptual learning, differentiation and acting so as to produce structured patterns of information specifying achievement of task goals (see section 2.4). Concurrent augmented feedback is different to feedback in the KR/KR tradition in that it enables *augmented perception* for the control of movement.

This suggests a novel interpretation of the guidance effect in the context of concurrent augmented feedback. If learning a task is characterised by education of attention towards task-relevant information, then it is possible that visual augmented feedback (such as on on-screen graph or display) encourages the development of perception-action strategies which are only effective in the presence of the feedback display. With visual feedback, learning some complex motor tasks can be sped up dramatically (from days down to a single session), at the price of a severe guidance effect (Kovacs & Shea, 2011). Instead of watching his/her limbs, the learner is typically encouraged to watch a graphical representation of the movements made by his/her limbs. The patterns of sensory stimulation available through interaction with a graphical display often differ dramatically from those available through interaction with the intrinsic features of the motor task. On post-acquisition no-feedback retention tests, performance may decline because participants have not learned to pick up and use kinematic information from the limbs; the education of attention developed during the acquisition stage is effectively wiped away. Interestingly, sonification as feedback does not necessarily require visual attention to be oriented away from task-intrinsic information, which could in some cases make sonification a more appropriate form of feedback than a visual display (Fitch & Kramer, 1994).

2.6.6. Feedback summary

Relationships between action and perception have not been the central focus of historical research on augmented feedback (in the KR/KP tradition). While J. A. Adams (1971) made clear that a “perceptual trace” (p. 123) was a necessary component in a full theory of motor skill learning, his conceptualisation was that of a memory store which held an ‘image’ of sensory feedback to be used as a reference for future trials – a knowledge structure to contextualise and help implement performance updating with KR. More recently, Thomas and Thomas (1994) have argued that traditional ‘knowledge-based’ approaches to motor skill learning underplay the role of selective sensitivity to perceptual information in expert performance[5]. Advances in high-speed data capture and display technologies (those which are today used to provide concurrent augmented feedback) enable the construction of new perception-action couplings, which allow the role of perception in motor skill learning to be studied more fully.

2.7 The role of auditory information in perception and action

While recent trends (increasing interest in augmented feedback delivered concurrently with movement) are bringing feedback and motor learning research more in line with

[5]Thomas and Thomas do acknowledge the importance of procedural and declarative knowledge in novice skill acquisition and performance (especially in light of Fitts and Posner's (1967)well-known 3-stage model of skill learning). In line with this, they argue that ‘knowledge-based’ approaches to studying feedback might only be applicable to the early ('Cognitive') stage of learning.

ecological concerns about the role of 'online' or 'live' perception, the focus remains primarily on the visual modality (Pizzera & Hohmann, 2015). The role of auditory perception in motor control and learning has historically been (and remains to some extent) understudied (Murgia & Galmonte, 2015). Despite this, the extant research indicates that auditory information can play a significant role in motor tasks, thus implying that extra action-coupled sound (i.e. movement sonification) could also be beneficial.

2.7.1 Natural action sounds in rhythmic tasks

Where perception-action research has considered auditory information, findings have revealed sound as a surprisingly important component of motor control, especially in tasks characterised by a strong rhythmic or temporal component. The value of rhythmic auditory information in such tasks is that it can provide a clear temporal structure with which motor behaviour can be synchronised. In relatively constrained sensorimotor timing (i.e. tapping) experiments, Repp and Penel (2002) show that small timing asynchronies in rhythmic stimuli are more readily adjusted for by participants' tapping behaviour when the stimuli are auditory rather than visual. In perceptual discrimination tasks, it has been repeatedly shown that participants can identify the sound of their own motor performance from a set of similar stimuli either produced from the movements of other participants, or synthesised. Flach, Knoblich and Prinz (2004) tested this with the recorded sounds of clapping, and found that participants were able to identify their own clapping patterns even when claps were replaced by generic sounds; the temporal patterning was still informative. Similar results have been shown in perceptual discrimination tests conducted with sound stimuli derived from various activities, including piano playing (Repp & Knoblich, 2004), golf (Murgia, Hohmann, Galmonte, Raab, & Agostini, 2012), and hurdling (Kennel, Hohmann, & Raab, 2014). Sound produced by one's own action seems to be readily identifiable as such. These studies and many others like them (see Pizzera & Hohmann, 2015, for a review) indicate that there are strong links between auditory perception and action, especially for actions performed by the listener. However, a stronger case for the usefulness of auditory augmented feedback can be made by considering more ecologically valid experimental tasks - i.e. those in which perception and action happen concurrently.

2.7.2 Relations between auditory perception and action

The role of sound in motor performance has been less frequently studied in experiments which use live sound, coupled to participant movement (outside of the tapping literature, which showcases the fundamental importance of sound for small-scale rhythmic synchronisation, see: Repp, 2005). However, several studies have investigated the extent to which action is dependent on task-intrinsic auditory feedback by experimentally manipulating the available sounds. Some motor tasks which are primarily sound-oriented (i.e. tasks in which motor performance is judged primarily on the basis of sound), such as speech and music-making are extremely susceptible to disruption via manipulation of auditory feedback. Howell, Powell and Khan (1983) manipulated the temporal profile of speech sounds fed back to speakers. They found that a short delay (200 ms) in feedback induced severe decrements in performance (stuttering, involuntary volume and speech rate fluctuations). The same disturbances were observed when the 'amplitude contour' of speech was distorted, without any temporal delay. The authors also performed an additional experiment with experts in Morse code - in which a delay in sound feedback again proved deleterious to performance. In music, Pfordresher (2009) reports that for expert pianists, outright removal of auditory feedback does not substantially affect motor performance, but that small temporal manipulations to feedback "can profoundly debilitate performance, to the extent that a skilled performer sounds like a beginner" (p. 183).

Even in some tasks which might appear not to be functionally dependent on auditory feedback, auditory distortion can affect performance. For example in straight locomotion, the primary perceptual outcome criterion can be expressed visually as a central expansion in optic flow (Lee, 1976); sound is incidental. However Menzer et al. (2010) show that manipulations to footstep sounds heard while walking (a variable temporal delay, up to 1800ms) have a systematic effect on gait period and walking speed; a longer delay led to more pronounced effects. Kennel et al. (2015) used a waist-mounted microphone and headphones to temporally disrupt the sound of footfalls for hurdlers (180 ms delay), which induced small changes to kinematic aspects of stepping. However, participants quickly learned to either compensate for delayed feedback or ignore the sounds altogether, as the effect disappeared after the first trial. Indeed, Pfordresher (2009) suggests that skilled performers in many domains may be able to quickly adjust to systematically altered auditory feedback, especially where it is not a primary functional outcome of the task.

2.7.3 Effects of skill in auditory-motor tasks

In line with arguments developed earlier in this chapter, listening may be conceptualised as a skill, characterised by a particular readiness to move a certain way and to pick up task-relevant information for use. Much as how chess grandmasters visually assess the state of play by performing different saccades between different features of the board than novices, expert listeners show the ability to efficiently differentiate useful sound information, in a way which is constrained by their practiced task. Arguably all the effects of sonic manipulation on motor performance reviewed so far in this section are evidence of learned styles of picking up sound in service of action; it follows that disruptions to auditory perception affect action because many actions are skilfully controlled (to some degree) on the basis of action-coupled sound. As a more direct example, Cesari, Camponogara, Papetti, Rocchesso and Fontana (2014) showed that only expert skateboarders were able to modulate their action appropriately in time to the recorded sound of skateboard jumps, whereas novices were unable to do so, being unfamiliar with the task as a whole. Skaters were able to pick out the relevant auditory information and put it to use in action coordination.

Pfordresher (2009) proposes that manipulations to the ‘content’ of musical feedback[6] (i.e. pitch rather than temporal onset) might affect pianists of different skill in different

[6] The theoretical approach of this thesis would not endorse a distinction between structure and content of information in this case; musical pitch variation can be conceived as perceptually differentiable variation in information structure, in accordance with the wine and paint examples explained in section 2.4.1.

ways. A systematic shift in the tones produced by notes on a keyboard (transposed six semitones higher) has little effect on motor performance of skilled pianists, but has a negative effect on piano novices. Pfordresher suggests that skilled pianists have learned to perceive melodies in a generalised, Gestalt fashion, regardless of transposition, i.e. that they do not perceive (as great) a conflict in the task relative to novices due to task relevant experience.

Effects of auditory-motor skill in one domain can sometimes transfer to another. For example, the meanings of many words in Mandarin are signalled by the pitch in which the word is spoken. Giuliano, Pfordresher, Stanley, Narayana and Wicha (2011) found that native speakers were much better at fine pitch discrimination tasks than non-speakers. Given that fine tonal differences in sound are behaviourally relevant for speakers of Mandarin on an everyday basis, it is likely that education of attention underpins this result (J. J. Gibson & E. J. Gibson, 1955). This finding could also be regarded as an example of a culturally-situated skill - fine tonal differences are differentiable by native speakers of Mandarin because the differences are meaningful in their practiced form of life (Rietveld & Kiverstein, 2014).

2.7.4 Altering tasks and task performance with sound

Watanabe and Shimojo (2001) show that the application of sound to a visual task can alter how the stimuli are perceived altogether. In a series of experiments, two black discs moved towards each other on a screen, passing through each other and stopping at each other's original starting location. With the addition of a short burst of sound (a 1800 Hz tone) at the moment of 'contact', participants reported that the discs now 'bounced' off each other rather than passing through, despite no change whatsoever in the visual stimulus. The extra auditory information had changed how the task as a whole was understood.

Ostensibly task-irrelevant sounds can have also an effect on motor performance in simple motor tasks, such as reaching to grasp objects. Castiello, Giordano, Begliomini, Ansuini and Grassi (2010) paired reaching actions with either a sound congruent with the material of the object being reached for, or an incongruent material sound. They found that reaching was facilitated by congruent sounds, in the form of faster movement speed and smaller grasping aperture, and conversely inhibited by incongruent sounds, relative to control. In a similar study which placed some demand on audition as a primary information source, required for correct task performance, Sedda, Monaco, Bottini and Goodale (2011) found that participants spontaneously adjusted their grasp aperture according to the size of the object heard to drop (one of two wooden blocks), whether vision was available or not.

These experiments show how moving individuals seem to automatically express implicit knowledge of learned action-sound relationships through altered styles of bodily movement. More simply, certain sound types imply a certain kind of task. Participants in these experiments alter their movement style in line with the task constraints implied by sound. On this basis it might be possible to encourage task-specific movement styles with the imposition of certain sound morphologies. To illustrate this idea, Rodger and Craig (2011) conducted an experiment in which participants were required to synchronise wide, planar hand movements to a sonic pacing stimulus. They found that the type of sound selected for the pacing stimulus had an effect on the kinematics of synchronisation movements produced by participants. Continuous pacing sounds (with an amplitude peak to mark synchronisation onset) encouraged more harmonic/sinusoidal synchronisation movements than discrete sounds, which were associated with more discrete movements. Despite participants only receiving instructions to keep their movements synchronised to the temporal interval, measurable differences in movement style emerged depending on the type of sound involved. This may reflect the expression of implicit, body-involving (or, ‘embodied’) knowledge of invariant sound-action relationships concerning the kinds of movement events usually specified by continuous vs. discrete sounds in everyday life (Rodger & Craig, 2014).

Different tasks are affected by manipulations of available auditory information in different ways. Furthermore, individuals who are skilled at the task in question often exhibit different reactions to novices. This indicates a degree of agent-task-specificity for the role of sound; it is likely that sonic-informational variables are task-critical inasmuch as they afford specific task-relevant uses, for an appropriately skilled agent (Steenson & Rodger, 2015). Even in tasks where sound is ostensibly extraneous, differences appear in task performance, as if the task had in some way, been altered.

2.7.5 Cross-modal interactions between auditory and visual/tactile perception

It may be a mistake to deal with auditory perception and action in isolation from other sensory modalities (as this section has to some extent). The idea that sensory information is cognized somehow differently depending on the medium in which it is detected is a widely-held assumption of Psychology. However, the assumption that, for the perceiver, there *are* discrete sensory systems is being questioned. Phenomenological theorists (and later, proponents of the Ecological approach to perception) have long held that events are perceived first in a holistic, multimodal manner. In everyday terms, splitting the event of one's baseball bat hitting a softball into the visual, auditory, haptic, and proprioceptive and motor systems is really an abstraction from the holistic experience of a swing and hit.

The 'supramodal brain theory' (Rosenblum, Dias, & Dorsi, 2016) suggests, based on a recent reinterpretation of a catalogue of cross-modal perception research, that the brain is largely "agnostic" about information modality. It suggests that the (what have traditionally been assumed to be functionally discrete) sensory cortices are not necessarily specialised to deal with information of a specific modality, but to perform certain kinds of tasks, making use of whatever information best supports performance in that context. Much of the evidence for this theory comes from fMRI research which shows cross-activation in cortical areas traditionally thought to deal with only a single modality of information - for example, the repurposing of the 'visual' cortex in the blind towards involvement in spatial cognition more generally (Cecchetti, Kupers, Ptito, Pietrini, & Ricciardi, 2016; see also: Kolarik, Cirstea, Pardhan, & Moore, 2014) and the recruitment of the 'auditory' cortex for sign-language communication in the deaf (Bavelier, Dye, & Hauser, 2006). In healthy subjects, tasks can be presented in a single modality and still elicit activation in multiple individual sensory cortices, most especially speech (Rosenblum, 2008). Similarly, behavioural performance on perceptual and motor tasks can often be maintained following a switch in modality. For example, lip-reading from one speaker conferred a performance boost at a secondary speech-in-noise auditory detection task which used that same person's voice (Rosenblum, Miller, & Sanchez, 2007). The same transfer effect also worked in reverse: listening to one speaker led to improved performance on a task requiring lip-reading from the same speaker (Sanchez, Dias, & Rosenblum, 2013). With the use of technologically-based sensing systems and wearable displays, one modality can be directly transformed into another and used to perform tasks. Bach-y-Rita & Kercel (2003) review several decades' worth of research showing that sensory substitution technologies (most often vision into haptics) can enable adaptive motor task performance after only a short period of practice. In some cases, participants even report 'visual-like' experiences with use of haptic substitution devices.

This framework could provide some physiological basis for some commonly-reported 'metaphorical' mappings between patterns in different sensory domains, including the 'boubah-kiki' phenomenon, in which shapes of a certain form (round and bulbous vs. spiky in this case) are reliably paired up with names which seem orally congruent (Ramachandran & Hubbard, 2001). Similar metaphorical mappings have been investigated in sound, e.g. the 'SMARC effect', in which stimulus-response compatibility is found between 'high' tones and the sense of something being 'high' in space (Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). A common sensorimotor task or skill could be the link between the domains of sound and movement. In a cross-cultural study between the USA and Cambodia, Sievers, Polansky, Casey and Wheatley (2013) make a convincing case for common dynamic structures perceivable through motion, form and music. Their participants produced consistent patterns of motion and form in a manipulable bouncing ball animation and accompanying manipulable music when given an emotion-based prompt, e.g. happy or sad. Klapp and Jagacinski (2011) argue that some 'perceptual Gestalts' seem not to be tied to specific sensory receptors, and indeed can be said to unfold over time - in vision, sound, and crucially for the present discussion, in the motor system. Very young infants for example, can recognise a pattern explored orally (a nub-covered pacifier, hidden from vision) and visually (a nub-covered display). These perceptual Gestalts can be said to have a constant pattern, regardless of the mode of presentation. Johnson (2007) argues that this constancy is in the pattern of stimulus intensity over time, regardless of the part of the nervous system which detects it.

The supramodal brain theory places most of the explanation for cross-modal mapping in the workings of the brain. However, some related arguments have arisen from Ecological Psychology. Stoffregen and Bardy (2001) argue that true specification of real-world dynamics often requires redundant information across multiple sensory modalities. Ecological Psychological theory has often referred to ambient sensory or energy 'arrays', conceptualised as structured manifolds of world-specifying information available within an ecological system (Jacobs & Michaels, 2007; Mossio & Taraborelli, 2008). In a thusly-defined ecological system, agents can interact with ambient arrays to generate information which is inherently spatiotemporal[7] and of a higher order than could exist without definition of both environment *and* agent activity (see section 2.2). Stoffregen and Bardy suggest that it is logically consistent with this approach to say that agents perceive events and objects by sampling a superordinate 'global array', which ontologically subsumes the separate visual, haptic and auditory arrays traditionally discussed.

> "*The possibility that specificity exists solely in the global array provides the possibility of direct perception, but only if the senses function as a single unit. To accept this possibility requires rejection of the assumption of separate senses. A view emerges in which perception consists not of a group of systems working in parallel (and often in conflict), but of a single system whose parts operate as a unit to pick up information that is available only to the unit.*" (Stoffregen & Bardy, 2001, p. 211)

---

[7] The addition of time as a dimension in the Ecological conception of 'information' is essentially what allows direct perception in the first place, through 1:1 mappings between dynamic events and resultant higher-order information (J. J. Gibson, 1966; Withagen & van Dijk, 2016).

Stoffregen and Bardy's argument is that informational redundancy is fundamental to perception. The information available through interaction with a single ambient energy array is not always enough to enable 1:1 specification, but the detection of higher-order specifying patterns across multiple modalities is possible in principle and can enable inherently multimodal direct perception.

2.8 Concluding remarks

It is reasonable based on the literature reviewed in the current chapter to take the view that dynamic interactions between brain, body and environment combine to produce a multimodal sensory experience of the world. It will be important to take forward the idea that perception of sound in a motor task is not neatly separable from concurrent pickup of information from other energy arrays. In fact, it is commensurate with Stoffregen and Bardy's approach to see the addition of movement-coupled sound to a motor task as a scaffold for access to higher-order informational variables - which could in principle be better-specifying information for perceiving and controlling motor performance.

While still somewhat speculative, these links between motional forms perceivable across sensory modalities, including sound, could be harnessed with the use of sonification as concurrent augmented feedback. If the pattern of stimulation between the required movements of a motor task and the sounds paired with movement can be brought into synchrony with each other (i.e. can be composed into a common perceptual Gestalt), it is conceivable that the result would be a kind of feedback which is perceptually integrated with the movements of the task in a way that more traditional forms of feedback are not (e.g. metrics, graphs, verbal feedback). Sonified feedback which perhaps draws attention to aspects of performance that should be corrected can be picked up whilst concurrently attending to all intrinsic sources of feedback, and may well integrate more readily with intrinsic sources, forming a more robust and persistent framework for learning. Certainly, there is substantial potential in the use of sound to guide movement coordination in otherwise silent tasks. The exact nature of sonification is something which needs to be considered very carefully, however. The relationship between the movement of the learner and the sound fed back is an area worth considerable attention, and is the subject of the following chapter.

Chapter 3

Sonification Mapping

Sonification, in the most general sense, is the use of non-speech sound to represent and convey information to a listener. However, “sonification” also refers to somewhat different things depending on the context of its use. In fact, the term can refer to a surprisingly wide variety of applications, research methodologies and theoretical positions, despite the relatively short time-span of public and academic interest in the general technique (Hermann et al., 2011).

The current thesis is interested in how digitally-generated sound might be used to inform a learner about their performance in a motor task, and guide him/her towards correct performance. For this application to be successful, it is crucial that the extra auditory information provided by sonification be understood by the learner. In considering this issue, I found that some of the most influential work and discussions pertinent to interaction with digital sound are effectively hidden on the other side of disciplinary boundaries. In this chapter I will review current trends and theoretical discussions in three fields which use sonification. I will start with current research in Psychology and related sciences, where sonification is used as augmented feedback for motor skill learning. This first section will more clearly explicate the aims of the current research project and thereby constrain discussion in the subsequent sections. I will then consider Auditory Display, the practice in which complex datasets are sonified so that listeners might experience their findings sonically, and perhaps gain new insights as a result. Finally, I will explore a subfield of

Music which is concerned with the design of new musical instruments which are engaging and require skill to use.

These three disciplines are making active use of sonification, driven by numerical data submitted to computerised sound synthesis. Although their aims are disparate (motor skill learning, data display and music-making respectively), all are, in my view, grappling with similar issues. Namely, the difficulty of producing sonic experiences which are informative and meaningful without a rulebook on how to do it, or any restrictions on how sound can be mapped to data. Each field has its own mapping problem. In Psychology, there is no consensus on how sound should be mapped to action to enhance motor control and learning, via augmented perception of movement. In Auditory Display, practitioners face the challenge of imparting knowledge to untrained listeners through carefully-designed soundscapes. In Music, researchers design instruments unconstrained by the physics of real-world sound production, and have the difficult task of explaining how skill and meaning can emerge in human-machine interaction. The aim of this chapter is to bring together insights from these three fields to inform a perception-action approach to the design of sonic-interactive motor tasks.

3.1 Sonification in Psychology

In recent years, sonification of movement has emerged as a viable method for the provision of augmented feedback in motor skill (re)learning. There is potential for sonification to be a highly-effective alternative or supplement to traditional methods of providing feedback in a sporting or rehabilitative context. Despite some experimental validation of its utility, controlled trials to test different methods of implementing sonification are still rare. A critical consideration which still needs both theoretical development and empirical investigation is the relationship between movement and sound: the mapping. The kinds of sound paired with movement and the way in which sound is modulated *by* movement have implications for the effectiveness of motor skill learning interventions. There are yet no accepted conventions for dealing with mapping, however some rough patterns are beginning to emerge from the experimental literature.

3.1.1 Mapping movement to sound

Sonification of movement entails the use of technology to generate live sound from human bodily motion (Höner, 2011). To achieve this, movement data is typically captured with the use of accelerometers, optical motion capture or force transducers and fed into a digital sound synthesis engine. Modern high-speed computing allows the corresponding sound to be produced with very little latency, so the user is effectively controlling live sound in real time with the movement of his/her body. Movement sonification has found application in sport, in which athletes can make use of sound information to more accurately time their actions (Kleiman-Weiner & Berger, 2006; Schaffert & Mattes, 2015; Stienstra, Overbeeke, & Wensveen, 2011). Additionally, therapeutic interventions have been designed involving sonification for rehabilitation of patients with motor disorders, in which patients use sound cues to supplement degraded proprioceptive feedback and to promote reacquisition of movement skills (Y. Chen et al., 2006; Maulucci & Eckhouse, 2001; Oscari, Secoli, Avanzini, Rosati, & Reinkensmeyer, 2012; Robertson et al., 2009; Rosati et al., 2013; Scholz, Rhode, Großbach, Rollnik, & Altenmüller, 2015).

In a recent review of concurrent augmented feedback presented in different sensory modalities, Sigrist, Rauter, Riener and Wolf (2013a) provide some evidence that auditory augmented feedback (movement sonification) might be as effective as feedback provided in the visual modality (which is more widely deployed and known to be effective under certain circumstances) in the right context. They report clear evidence that sonification could be effective in simple tasks (i.e. very reduced tasks which employ only a single effector, or few degrees-of-freedom, such as one-handed reaching/force production), but little direct evidence for efficacy in more complex tasks (i.e. with involvement of multiple degrees-of-freedom and competing task requirements). Direct comparisons between modalities of feedback in complex motor skill learning contexts, especially incorporating sonification, are rare. Sigrist, Rauter, Riener and Wolf (2013b) examined task learning with concurrent augmented feedback presented in three separate modalities: visual, auditory and haptic. Participants were required to practice a rowing activity in a simulator with feedback provided on alternate trials i.e. on 50% of the total number of trials. In the auditory feedback condition, *movement error* was sonified, i.e. sonic variations were produced using measured deviation from the ideal movement profile - on both the horizontal and vertical plane, as well as rotational timing deviation. Visual feedback was provided on a screen to the side of the participant, showing the target oar trajectory with live performance superimposed on top. Haptic feedback was provided via robotic manipulation of the handheld oar - physically guiding the learner towards the target trajectory. The difference in movement error between feedback trials and non-feedback was very noticeable for the visual group. When feedback was present, participants showed very low error compared to when it was absent. A similar effect, although less pronounced, was observed in the haptic group. On retention trials, performance by these two groups was significantly worse than on earlier feedback trials (although some degree of learning relative to initial baseline was evident).

Unlike the groups practicing with visual and haptic feedback, average performance in the auditory group did not vary based on the immediate presence of feedback. Performance in this condition was highly variable between individuals and seemed to be entirely unrelated to the availability of augmented feedback information. This is an unusual pattern of performance to see in an augmented feedback experiment; scores in the visual and haptic conditions were much more typical - in that performance was improved in the presence of feedback and less so in its absence (Maslovat, Brunke, Chua, & Franks, 2009; J. H. Park, Shea, & Wright, 2000). An overall effect of learning (i.e. improvement in no-feedback performance from baseline) was not actually found in the sonification condition. One important thing to note when interpreting these results is that the authors acknowledge that their sonification prototype might not have been fit for purpose. Several participants reported great difficulty extracting the relevant performance information from their feedback. This experiment underscores the importance of effective sonification mapping design. If participants cannot extract meaningful, performance-relevant information from what they hear, learning might actually be hindered rather than enhanced. The next several sections will consider the various ways researchers have designed movement sonification to avoid this problem, with particular attention directed to the mappings used.

3.1.2 The mapping problem in motor skill learning

Given that sonification is generated digitally, there are few real constraints on how movement might be mapped to sound. Rather, there exists a real problem of *too much choice*, and with the lack of established guidelines in the field to narrow down these choices to an acceptable range, there is a risk that mapping decisions could be made which are less useful, or even detrimental to learning. Effenberg (2005) correctly asserts that, "An almost endless amount of options are available to transform data into sound" (p. 1). In the case of Sigrist et al. (2013b) above, it is theoretically possible that a workable sonification mapping might yet be found which could enhance motor skill learning in the rowing task. There are several main styles or types of sonification which are differentiable based on the structure of the information they provide to a learner. In the next section, some of the more frequently-used mapping types will be elaborated, with some examples of their implementation.

3.1.3 Error sonification

Error sonification is one of the most commonly-used solutions for human movement sonification in a motor skill learning scenario. The technique assumes that there is an exemplary movement profile or precise target movement for the to-be-learned skill. The movement variable which is sonified here is deviation from the ideal movement profile. Note that this information is not direct sonification of movement by itself, but indirect and abstracted sonification, in that it describes movement *relative* to a criterion. In principle, the learner should be able to use the extra auditory information generated by their movements to understand how they have deviated from ideal performance, and use that knowledge to correct the error immediately or on the next practice trial. This is the same style of sonification employed by Sigrist et al. (2011; 2013) in a rowing task.

A successful implementation of error sonification is presented by Konttinen, Mononen, Viitasalo and Mets (2004). Army recruits practiced rifle use at a shooting range and had their performance error sonified. In contrast to Sigrist et al., this error sonification was concerned with only one dimension of error - the deviation in position of the gun barrel from the target bullseye. As recruits moved their aim closer to the target, the pitch of a sine tone increased to a maximum at the centre, and vice versa to the edge of the target. The sonification group had their aiming sonified alternately on 50% of trials. The primary performance measure was shooting score, i.e. shot accuracy. There was a slight benefit of sonification during acquisition relative to the control condition (who practised with only knowledge-of-results feedback about shot accuracy). Interestingly, that slight benefit of sonification was observed even on trials in which sonification was withheld, which indicates that the recruits were not simply using sonification as a guide to find the bullseye, like a beachcomber might use a metal-detector. The sonification group additionally showed improved performance relative to control in delayed no-feedback retention tests, suggesting that sonification had a lasting benefit. This study might stand as evidence that more simple error sonification (i.e. sonification of fewer dimensions of error, as compared to the rowing example of Sigrist et al.) is more effective in biomechanically complex tasks; however, there is an issue with comparison between these two cases. Where Sigrist et al. intended to encourage the production of a pre-set movement trajectory within an acceptable range, Konttinen et al. were interested in enhancing perception of rifle stability, a crucial determinant of shooting success. Due to the mapping, shaky barrel movements were accompanied by wavering pitch, and steady aiming produced smooth changes in pitch. Learners could use this information to learn to hold the rifle more steadily, which, it is proposed[8], led to improvements in the outcome measure: accuracy. Although this is a clear example of error sonification in that deviation from a target position (i.e. accuracy) was sonified, the outcome was not functionally dependent on sound – and performance may actually have been enhanced in a less-direct manner.

Other motor learning studies have investigated error sonification in the task of reaching for a target and motor adaptation to systematically altered conditions. In this task, participants are required to reach for a stationary target a fixed distance away, typically without vision of their hand, relying instead on the artificial feedback provided by the system. These experiments share some methodological similarity with classic KR experiments, in which the hand is hidden and the learner must make trial-by-trial corrections to performance using verbal feedback (see section 2.6.3). Oscari et al. (2012) mapped positional error (relative to the target) on the *x* axis (i.e. left-right reaching error) to the amplitude of a pink noise generator so that greater positional error during a reach produced louder noise. Additionally, error relative to the target was mapped to a stereo-panning function, which mapped the pink noise more into the left ear when positional error was negative (i.e. left of the target) and into the right ear when error was positive. Targets were presented at a single location and participants learned to use the auditory feedback to guide their movements to the target over the course of a practice stage. The authors found that participants were able to use the auditory information to move correctly during a force-field perturbation phase and exhibited corresponding after-effects in motor performance whether they had been presented with sonification of error, or more direct visual information. In a similar study, Schmitz and Bock (2014) mapped a the pitch of a sine tone to horizontal deviation from target position so that pitch increased with error in either direction. Perfectly-targeted reaching produced a tone of 1337 Hz. Error from the target was additionally mapped to stereo-panning, the same way as described for Oscari et al. (2012). In this task however, the position of the targets changed between six possible locations. Participants

[8] Rifle kinematic data were not reported by Konttinen et al. - making this claim speculative, but plausible based on performance data in the acquisition stage.

used the target-relative sonification mapping to guide themselves to each target as quickly as possible. Again, following adaptation to systematically altered feedback, after-effects were observed in the sonification condition, similar (but not as severe as) in another condition, in which direct visual feedback was provided.

A subtype of error sonification is ‘alarm sonification’, in which deviations from a prescribed ideal are signalled by a sound meant to alert the learner. The difference between alarm and error sonification is in the nature of the mapping; alarm sonification is binary - either on or off. The degree of error for which the alert activates is typically set by the sonification designer in line with the requirements of the task. Baudry, Leroy, Thouvarecq and Choller (2006) sonified the performance of gymnasts practicing on a pommel horse. Their system incorporated a device in two parts, with sensors attached to the top of the back and to the back of the knee. One of the markers of good performance on the pommel horse is straightness of form between the legs and upper body. Flexion between the two is undesirable. When the measured angle between these two segments reached <20°, a buzzer activated to alert the participant to correct their form. The advantage of this kind of sonification is that it is very easy to use, assuming it is employed in a task in which the required correction is obvious. The authors found improved performance relative to control by the end of a practice phase with the sonification and the same pattern of results on a retention test two weeks later.

It is very uncommon for different kinds of sonification to be compared to each other in a controlled test of motor skill learning. One such example comes from Rosati, Oscari, Spagnol, Avanzini and Masiero (2012), who implemented three kinds of sonification in a movement-tracking task. This required learners to track the position of an on-screen target, through a handheld joystick. Two versions of the task involved the sonification of positional error between target position and the position of an on-screen dot, controlled through the joystick. Position error was measured continually and mapped separately to either the pitch or amplitude of sound generated via the technique of formant synthesis (which sounds like a human voice making a vowel-sound). These mappings did not lead to improved performance by the end of a practice phase relative to a control condition which practiced with only the on-screen visual information. The authors additionally implemented a more direct mapping between movement and sound, in which the movement of the to-be-tracked dot was mapped to the sound of a rolling ball to augment perception of movement velocity. This version of the task did produce improved performance relative to control. Boyer, Bevilacqua, Susini and Hanneton (2016) performed a similar study, replicating the tracking task and some of the conditions from Rosati et al. (2012). They compared positional error sonification, sonification of target velocity and sonification of pointer velocity (which was participant-controlled), by mapping the value of each variable to the centre frequency of filtered, continuous white noise. In all conditions, the centre frequency was scaled between 80-4000 Hz depending on performance. This study reported a mild guidance effect of error and target velocity sonification, but not of pointer velocity sonification.

3.1.4 Statement on error sonification

These results present a mixed statement on the efficacy of error sonification. It is possible that this inconsistency is not a reflection of the style of sonification itself, rather its suitability with certain types of task and even certain ways of defining learning. For instance, sonification has been shown to be efficacious in altered-feedback reaching tasks, showing that participants *can* make use of error quantities provided through sound (Oscari et al., 2012; Schmitz & Bock, 2014). This task is encapsulated within a single action: a reach with a clearly-defined trajectory. Every practice trial starts from an identical position and the feedback mapping reports veridical information describing each reach relative to the target trajectory. In such a constrained environment (sound and proprioception are the only feedback available) with a task which unfolds over a short timescale, a strategy of trial-to-trial corrections is possible. These tasks are set up such that a learner can use feedback as knowledge-of-results to update a motor plan, similar to participants in classic augmented feedback experiments who could only use proprioception and basic KR to tell if they were making the right movement (Adams, 1971; see also section 2.6.3). It remains an open empirical question whether error sonification is appropriate in more continuous movement tasks, but the extant literature hints that it might not be. Rosati et al. (2012) report no benefit of error sonification in a continuous tracking task; Sigrist et al. (2013b) find the same in rowing.

To elaborate further, recall that error sonification is an abstraction, or description of movement performance relative to the target, in sound. To understand and make use of this information requires some elaboration of incoming information in light of a remembered mapping rule. It may be that very simple, repetitive tasks (such as reaching) are more suitable for error sonification precisely because their repetitive nature allows participants to use sonification as an intellectual marker of performance, akin to knowledge-of-results. More continuous, freeform tasks, in which conditions change moment-to-moment, may be too attentionally demanding to allow for concurrent extraction of an error value from sound. The exception to this proposed rule is Baudry et al. (2006), in which the task was extremely complex (pommel horse), but the mapping simpler (alarm rather than continuous error) and the required correction always obvious (straightening form). Konttinen et al. (2004) represents a special case, as although undeniably error sonification, participants in that study seem not to have used sonification to improve their aim, rather to perceive and better control rifle stability.

3.1.5 Direct sonification

An alternative solution for action-sound mapping is to design sonification which is controlled by some quality of movement itself, rather than the movement relative to a goal trajectory, as in error sonification. This category is extremely broad, as there are about as many ways to measure movement as there are sounds to map to. Most often for sonified tasks which employ a direct style of sonification, the position (or a derivative of position) of the end-effecter (e.g. the fingertip in a pointing task) is tracked and sonified. The aim is that the learner will be able to integrate the extra sound information with task-intrinsic sources of information (vision, proprioception, haptic information), and perceive his/her movement performance more finely as a result of information synergy (De Gelder & Bertelson, 2003; Stoffregen & Bardy, 2001).

Boyer et al. (2013) report an example of direct hand sonification used in a pointing task. Participants were asked to point at audio targets, presented through headphones, spatialised binaurally in 3D using head-related transfer functions[9] and updated live as participants turned their head. The position of the hand was sonified with white noise and spatialised using the same technique. In this way, a participant could hear where his/her hand was in 3D space as if it was emitting sound. The authors did not find that pointing accuracy was any better when participants could hear the position of their hand compared to when they could not. Additionally, perturbing the auditory 'location' of the hand position on some trials did not produce compensatory effects on performance as expected. The authors speculate that participants may have experienced confusion between the sound of the target and the sound of the hand (both were the same, white noise), or alternatively, that spatialised audio, tied to the position of the hand, is a perceptual variable to which participants are not sensitive enough to use effectively. They suggest that sonification of another parameter, perhaps related to kinematics, might be more effective.

Movement kinematic variables are frequently used for sonification in studies on movement rehabilitation. Practice of simple, everyday movements are the cornerstone of stroke rehabilitation, most commonly, reaching to grasp an object (Dobkin, 2004). Movement after stroke is often impaired; arm movements often lack the sinusoidal nature which is characteristic of healthy individuals, and some sonification interventions have been designed to assist in recovery of this function (for a review, see Rosati et al., 2013). Additionally, sonification of movement can stand in for proprioceptive feedback, which is often degraded following stroke. Wallis et al. (2007) designed an immersive audio-visual system to retrain smooth reaching actions. The task involved repeated reaches and

[9] A head-related transfer function (HRTF) describes how a sound signal emanating from a defined location is picked up by each ear, accounting for the distance between ears and the physical properties of the head, shoulders and body, which modulate the signal. HRTFs can enable the synthesis of convincing 3-dimensional audio stimuli.

withdrawals of the arm, each of which was mapped to a chord progression. Participants needed to complete the reach, otherwise the corresponding chord progression would have no "choral resolution" (p. 4). As the task continued, different chord progressions were played, and more instruments became audible as various performance variables improved. As the authors describe it:

> "*Certain instruments are linked to aspects of the motion. The 'orchestra' instrument, for example, which is usually the first background instrument to be included, is linked to elbow openness. Initially inaudible, the volume of the orchestra swells as the elbow angle increases. This instrument mapping is motivated by those stroke patients whose opening and flow problems stem from reduced ability to extend their elbow*." (Wallis et al., 2007, p. 3)

Although a small-scale study (three stroke patients), the results were promising, showing an improvement in several measures of reaching performance. In a similar study, Scholz et al. (2015) sonified hand movement in a 3D cube-shaped zone in front of participants with stroke impairments. Musical tones were played as participants moved within the cube. Vertical position was mapped to pitch; higher position produced higher-pitched tones within the key of C Major. Movement from left to right altered the 'brightness' of the tones, becoming brighter to the right of the cube. Training in this system was structured more like a music lesson than a traditional movement training regime; participants practised several bespoke exercises, including learning to play simple folk melodies inside the cube. Results showed promising improvements in both tests of motor function and emotional wellbeing in the two experimental patients. No improvements were observed in two control patients, who underwent a similar regime without sound.

Other examples of the direct style of sonification have been reported by Effenberg and colleagues. Effenberg (2005) describes a procedure in which vertical downward force during jumping was mapped directly to pitch and amplitude of a synthesised formant so that amplitude and pitch increased with jump intensity. Observers of these sonified jumps were better able to perceive relative height when asked to pick between two examples.

Furthermore, participants were able to reproduce heard jumps when their own jumps were sonified using the same parameters. Vinken et al. (2013) mapped velocity of limb movement to sound amplitude, left-right movement to stereo balance, front-to-back movement to spectral patterning and vertical movement to pitch of a synth sound and found that participants were able to reliably match heard sounds with the everyday actions they sonified (e.g. stirring a pot, brushing teeth, filing nails) despite not being instructed as to the nature of the mapping. Effenberg, Fehse, Schmitz, Krueger and Mechling (2016) sonified grip force, footrest force, length of pull (mapped to pitch and amplitude of a continuous sound) and seat position (maximum and minimum position was sonified on an 'event' basis) in a rowing machine task, in which the goal was to match the force and kinematic profiles of an exemplar. Participants who practiced with sonification performed more like the exemplar than did those in a control condition throughout practice, and maintained this good performance into no-feedback retention. The authors were careful to ensure that perceptually-salient qualities of sound were matched by the quality of intrinsic feedback, by setting lower limits on the force required to activate sound modulation. They state:

> "*...forces could only be acoustically perceived when they were also kinesthetically clearly perceivable*" (Effenberg et al., 2016, p. 6).

Some researchers have used sonification in a manner not dissimilar to practice with a simple musical instrument, with task acquisition instantiated as repeated attempts to match a 'goal' sound. It has been shown that this style of direct sonification is particularly effective in bimanual out-of-phase rhythmic movement tasks where visual augmented feedback is typically employed (Heitger et al., 2012; Ronsse et al., 2011). The ideal sound profile is played periodically throughout practice (in these cases, a two-tone galloping rhythm) with participants attempting to flex and extend their hands at the wrist in a pattern which will produce the same pattern of sound. After removal of live sonification, good performance remained. In a one-handed, four-finger rhythmic task, van Vugt and Tillmann (2015) observed benefits of sonification tied to key-presses. Each key press was sonified with a synthesised tone played on a wood-block instrument. Benefits of sonification again remained after the removal of live sound.

Some of the most sonically complex and aesthetically interesting sonification in the context of motor skill learning comes from Danna and colleagues (Danna et al., 2014; Danna, Fontaine, et al., 2015; Danna, Paz-Villagrán, et al., 2015). They sonified handwriting kinematics in a task which required participants to learn to draw novel symbols with their non-dominant hand. The aim was to improve handwriting fluency, in particular to increase movement velocity and smoothness. Their mapping used intuitive physical modelling which produced a smooth rubbing sound from continuous pen movement, squeaking when velocity was too slow and cracking when the pen stopped moving. They found kinematic performance enhancement effects of practice with the system after a period of use. Additionally, listeners could discern the sound of someone who was writing in a smooth, fluid manner from someone who was having difficulty. The authors suggest that their system could be used to aid both handwriting training and the diagnosis of graphomotor disorders.

3.1.6 Statement on direct sonification

Considering implementations of 'direct' movement sonification, it is notable that few experiments report problems with participant understanding of the mapping. Indeed familiarisation sessions (during which participants explore the interface and learn how sound responds to movement), where they are reported, are typically short in duration. Vinken et al. (2013) even demonstrate that participants who have no familiarisation time can instantly match actions with their sonifications at much greater than chance level. Unlike error sonification, the direct style generally preserves the structure of information available from intrinsic sources, i.e. relevant events are both seen/felt *and* heard, with stimulation intensity usually roughly equivalent between sources over time. This structural equivalence between information in the auditory and intrinsic sensory arrays may enable perception of the sonified task as a coherent Gestalt (see section 2.7.5, for more on this idea), with sound functioning as a redundant scaffold for augmented perception of movement. There is also no need to consciously elaborate direct sonification to extract a coded score as in error sonification, perhaps making the direct style more suitable for complex motor tasks with increased attentional demands. Effenberg et al. (2016), on their mapping philosophy, argue along similar lines:

> "...*the processing of our kind of movement acoustics is not dependent on conscious cognitive processing, because the processing - even multisensory integration - is mandatory if the stimulus is hearable and certain criteria of intermodal convergence are fulfilled*." (p. 5)

It is reasonable to expect that direct sonification which is structurally coherent with respect to intrinsic information should produce sonified tasks which are relatively easy for a novice user of the system to understand. Furthermore, there is emerging evidence to suggest that learning novel tasks with direct sonification to guide performance may not lead to a guidance effect, i.e. feedback-boosted motor performance may not decline when direct sonification is withdrawn (Effenberg et al., 2016; Ronsse et al., 2011; van Vugt & Tillmann, 2015). Again, this may be attributable to the structural coherence between auditory augmented perceptual information and intrinsic information. Although sound is not an intrinsic part of the underlying motor task, perception of sound during practice does not entail perception of a wholly different sensory pattern. Compare direct sonification to, for example, certain forms of visual augmented feedback (e.g. graphical displays), which present a transformed perceptual pattern and lead to a guidance effect, (see section 2.6.5).

3.1.7 State of the art in Psychology

Sonification is becoming more widely adopted in Psychology where researchers aim to enhance motor performance and skill learning. Traditionally, research on the effects of augmented feedback have been concerned with only the visual modality, symbolic information in the form of numbers/graphs and guidance from a human coach (Magill, 2011). This is changing quickly. However, a challenge for the field going forward is the need for structured and inclusive discussion on sonification mapping. Rules and recommendations which might assist in the design of a sonification procedure for a new motor task generally do not exist. This makes it difficult to decide what an appropriate mapping for a given task or skill might be.

Researchers often intuit their own bespoke mappings for their task of interest and rarely provide theoretical justification or reasoning for the choice. In some cases, mapping decisions are transparently arbitrary (for example, mapping perfect reaching performance to 1337 Hz in Schmitz & Bock, 2014, which is a reference to 90s hacker culture, not a perceptually relevant frequency value!). There are vast differences in how tasks are sonified, both in terms of variable selection, and sound selection.

With regard to variable selection, some sonification systems are designed to guide the learner towards 'more correct' performance *from the perspective of the experimenter*. To elaborate with a hypothetical example, positional error may be identified as the variable of interest for analysis of performance data; decrease in this variable over time represents improvement on the task and the acquisition of motor skill. However, this is not necessarily an appropriate variable for sonification, as it may not be perceptually relevant for the learner. Task performance is often not explicitly considered from the perspective of the learner as a perceiver. Although in some cases, information variables which might be relevant to the user are identified by a task analysis, and these are sonified, with the aim of enhancing information pickup (Danna, Fontaine, et al., 2015; Effenberg, 2005), this is not common.

The use of interesting sound for sonification in Psychology is very much lacking compared to other fields which use the technique. By far the most common sonification solution is to select a measured variable and map it to the pitch of a sine tone (Dubus & Bresin, 2013), without regard to the fact that such a sound is challenging to listen to for any length of time (with possible consequences for motivation and therefore, performance), and also lacking in worldly meaning (see: Henkelmann, 2007). With a few notable exceptions (Danna, Fontaine, et al., 2015; Scholz et al., 2015), sonification in Psychology employs aesthetically impoverished sound and basic sound-synthesis techniques (noise, pitch-mapping, stereo-panning, etc.) which leave a vast sea of potential meaning untapped. Sound is a rich medium for aesthetic experiences, but it is not optimally harnessed in many cases (see Roddy & Furlong, 2014).

Reviewing published literature makes it difficult (at this point) to assess the suitability of sonification as concurrent augmented feedback for a novel task or skill to be trained. If a given experiment (such as Sigrist et al. 2013, which sonified multidimensional error in a rowing task) shows that learners perform worse under sonification conditions than when using feedback in other modalities, it does not tell us that sonification is less effective than these other options as a general rule. It does not even tell us that sonification is less effective or appropriate to use as feedback for this particular task. A different mapping could hypothetically be designed which engages the learner, provides the relevant perceptual information and enhances motor skill learning as a result. At this point, there exists no established means to tell a bad mapping from a good mapping, other than by experimental comparisons between individual mapping designs. Furthermore, the impact of a slightly altered mapping on functional performance in the kinds of tasks used in the motor skill learning literature may be quite subtle, meaning that behavioural experiments are likely to be underpowered to detect differences (e.g. Boyer et al., 2016). To return to an earlier point, the lack of discussion on mapping itself could be slowing progress in the field. Such a discussion would need to go well beyond a single experiment, or series of experiments, and if possible, branch out beyond the narrow subfield of motor skill learning. This discussion will be attempted in the following sections.

3.2 Sonification in Auditory Display

Sonification does not necessarily have to happen 'live' as in human movement sonification for motor skill learning. Straebel (2010) posits that sonification is essentially a metaphor for the relationship between sound and some measurement of the world. Throughout history there have been many acoustic compositions which have employed some translation from ‘extra-musical’ natural phenomena to the notes played by a performer with an instrument (Barrass, 2012). For example, the work of experimental musician John Cage frequently draws on the extra-musical, from the shape of stones in a Zen garden (which become graphically-demarcated glissandi on the score) to the positions of the stars (Cage, 1996). Straebel sees continuity between the romantic ideal of capturing the natural world in artistic expression, this kind of experimental music and modern auditory-display sonification. Indeed, sonification in auditory display can be seen as one manifestation of a tendency in modern musical composition which sees artists surrender part of their own decision-making to some non-human system (Supper, 2014; Willcock, 2006).

Straebel's general metaphor for sonification becomes auditory display in the context of a composer who intends for the listener to understand the underlying extra-musical information in its own right. The field of Auditory Display (as represented by the proceedings of the annual International Conference on Auditory Display – *ICAD*, which is in its 23rd year as of 2017) is therefore predominantly interested in communication and listener interpretation of information represented in sound. The most widely-accepted definition of sonification in auditory display comes from Kramer et al. (1999):

> "*Sonification is defined as the use of nonspeech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation*." (p. 3)

Sonification is often conceptualised as a means to experience a dataset sonically, rather than through more traditional means, such as visualisation. Given the sensitivity of the human auditory system to rhythmic/temporal information encoded in sound (Shea, Wulf, Park, & Gaunt, 2001), in some cases it might actually be more appropriate to hear a dataset rather than inspect a plot - for example if a dataset contains regularly repeating patterns, a disturbance in the pattern might be more readily detected by a listener than viewer. Speeth (1961) for example, found that when listening to audified seismic data, listeners could detect rhythmic differences between tremors caused by earthquakes and those caused by nuclear detonation, and also categorise each effectively. In neurophysiology, live sonification of neuronal firing is used to detect cells of interest. Indeed, the discovery of 'mirror neurons' was facilitated in this way (Gentilucci et al., 1988). Throughout the short history of *ICAD,* a wide variety of datasets have been subjected to auditory display, from the very large-scale to the very small, including: plant growth and animal migration patterns (Ballora, 2016), seismic data (McGee & Rogers, 2016), weather records (Flowers & Grafel, 2002), human electroencephalography (Baier & Hermann, 2008, as cited in Supper, 2014) and the interactions of subatomic particles (Sturm, 2000). As with visualisation, the promise of auditory display is knowledge about some natural phenomena which might otherwise happen beyond our ability to notice.

While Auditory Display and sonification for motor skill learning might seem relatively distant from each other, practitioners in both domains have a somewhat similar remit. From the perspective of a sound designer, sonification mapping in motor skill learning and auditory display are identical challenges; continuous, sometimes multivariate datasets are transformed into sound so that a listener might better understand events which would not normally produce sound. Listening to either kind of sonification entails a perceptual engagement with an auditory task. In many ways however, designing auditory displays for listener comprehension is the greater challenge. In listening to the sound produced by one's own actions - even if the mapping is complex - sonification is situationally-bound and possible interpretations of sound are constrained by the movements made by the listener. Furthermore, movement sonification allows for intentional exploration of the mapping through movement and feedback. In Auditory Display, sonifications are typically listened to without accompanying bodily motion. Dealing ostensibly with *only sound* as an information source has led researchers in this field to consider the semiotics of artificial sound in much greater detail than is usually attempted in a motor skill learning context. As such, literature in the field of Auditory Display contains valuable insights about how to create meaningful interactions with sound which can be applied in Psychology, where auditory aesthetics are infrequently considered (see section 3.1.7). In the following sections, critical discussions relevant to auditory meaning-making will be reviewed, with particular focus on insights which might enhance human movement sonification through the development of more meaningful interactions.

3.2.1 The problem of mapping complex data to sound

Auditory display is a rather young field whose influence outside academic circles is still small, although growing. Indeed, an persistent theme of *ICAD* is building the profile of auditory display and stimulating public interest (Barrass, 2012; St Pierre & Droumeva, 2016). A major hurdle faced by the field however, lies in the lack of established rules or theory for mapping data to sound for intuitive listener comprehension. Considered as a counterpart to data visualisation, the problem for auditory display is clearer. The rules for data visualisation and interpretation are well-established and understood by appropriately educated audiences, such that for example: separate lines on an area bounded by a set of axes are understood to represent separate variables, a segment of a circle represents a corresponding segment of the total dataset, and the height of a bar represents the value of a variable (Beniger & Robyn, 1978). The conventions of data visualisation as a practice have effectively spawned a syntax which enables an audience to use visual information to gain some knowledge about the displayed dataset. No such widespread or specific syntax exists for auditory display. Roddy and Furlong (2014) suggest that without an established syntax to relate the display to the data, Auditory Display practitioners run the risk of using arbitrary mappings and rendering datasets unintelligible.

Methods for transforming data to sound are effectively limitless, however for the sake of brevity I will deal mostly with issues pertaining to the technique of parameter-mapping sonification, which is the prototypical auditory display solution of mapping data variables to sonic variables[10]. Parameter-mapping sonification is ideally suited (in theory) to the display of multiple streams of meaningful information simultaneously, and such is very frequently attempted. For example, Ballora (2016) mapped data pertaining to Musk Ox and Caribou

[10]To be distinguished from other methods of auditory display such as auditory icons/'earcons', alarms and 'audification' (presenting a data waveform directly as a sound waveform). For a review of these and other techniques, see Dubus and Bresin (2013).

migratory patterns, as well as plant growth to different sonic variables, which were played back simultaneously to illustrate changes through the years. St Pierre and Droumeva (2016) adopted the same strategy for airborne pollutants in major Canadian cities - each pollutant was mapped to a separate sonic variable and played together to represent the overall state of air quality over several months. The advantage of multidimensional sonification lies in its flexibility. Mappings can easily be altered to suit the composer/scientist's aesthetic preferences or communicative purposes, and to facilitate comprehension by listeners. Worrall (2010) notes, however, that the potential of multidimensional parameter-mapping sonification is rarely, if ever, fully realised:

> "*The main limitation of [parameter-mapping sonification] is co-dependence, or lack of orthogonality (linear independence) in the psychophysical parameter space. Linear changes in one domain produce non-linear auditory effects... These perceptual parameter interactions can produce auditory artifacts that obscure data relations and confuse the listener*." (p. 2)

Parameter-mapped sonifications are typically designed for the express purpose of communicating something about the dataset itself, and it is often assumed that novel insights about the dataset and relations between variables will emerge by virtue of sonic presentation (Flowers, 2005). However, when listened to, perception of one sound-data stream often impinges on or in some way alters perception of other streams, as described by Worrall. This may produce several possible experiences for the listener: 1) unexpected perceptual overlap and obfuscation of one or more information streams, 2) perceived grouping (where discrete information streams appear to form coherent whole sounds unexpectedly different in experience from the sum of their parts) or 3) distraction from one sonic variable by more attentionally-salient other (due to unexpected factors like musical skill, early learning, cultural relevance of a certain timbre, etc.). The potential variability in how a given sonification might be experienced is enormous, even if general principles of psychoacoustic grouping are adhered to.

3.2.2 Contextualising the listening task

Perceiving patterns in a sonification which are maximally relevant/important for the task at hand can be a challenge, whether the task is to locate a physical target in the proximal environment or to find out when temperatures hit a historic low during the last ice age. It has been advanced earlier in this thesis that understanding a situation is the result of 'education of attention', or learning to pick up maximally relevant informational variables in the sensory array (see section 2.4.1). The same kind of task analysis can be applied to sonification. It may not be immediately obvious which feature or pattern in the available sound is most relevant to the task, but with careful sound design, it might be possible to guide listeners towards it.

Picking out, attending to and extracting the relevant information in a multidimensional stream of sound can be difficult for inexperienced listeners, as much with the use of an auditory display as a sonification system for motor skill learning (Sigrist et al., 2013b; Vickers, 2012). With auditory displays, some of this difficulty comes from the nature of the medium as a detached mode of data presentation which is abstracted away from its real-world source. Given the lack of a universal syntax, sonic parameters are symbols with no grounding, like letters in an unknown language (Roddy & Furlong, 2014). In recognition of this, sound designers in Auditory Display have created some techniques which can aid the listener in picking out the right information from a sonification. Take 'time' for example. Events unfold over time, and data collected about events is inherently temporal. The rules of data visualisation usually have ‘time’ represented as a display dimension, often as the *x* axis on a graph. When changes in a variable over time are to be represented as sound, it might make sense to preserve the temporal structure of the real event to facilitate listener comprehension. However, this is not always practical. Most datasets submitted to sound synthesis for auditory display are not captured from real-time events at a suitable temporal resolution to allow for this. The timescale of some datasets is in months or years (or even hundreds of thousands of years, e.g. Ballora & Kenney, 2014). Playing these datasets in real time is not realistic, so time is compressed to make listening to the whole sonification possible in a single session. The extent of this compression is necessarily variable between datasets and mappings, which may leave the listener in some doubt over the temporality of what they are supposed to be hearing about.

Simple solutions to this problem have been proposed, among which the most common of which is an auditory *x* axis, in the form of a regularly-spaced sound which can be taken as the passage of a temporal interval in the dataset. The sonification of Ox and Caribou migration patterns by Ballora (2016) is typical in this regard; as the beginning of each day is represented by a click, with an emphasised click on every fifth day. Smith and Walker (2002) suggest that contextual cues which do not themselves provide information about the dataset can nonetheless ease comprehension by better structuring the task for a listener (consider the difficulty of extracting useful information from a line graph without axes, for comparison). 'Time' is but one variable which might be relevant to comprehension of a sonified dataset. Scaling of other dimensions can be just as difficult to parse without additional contextual cues. Absolute share prices for example might be impossible to discern without an auditory cue to signal that a certain low or high-point has been passed - analogous to a *y* axis. Smith and Walker used dynamic reference tones which signalled each new low and high point for share values and concluded, cautiously, that listener judgements about share values were improved with this extra information.

Although extra sonic information can structure the listening task by providing context, there is, again, no standard of practice for its use. One sound designer might opt to use clicks to signal the passage of days, while another might prefer to signal hours, or use another kind of sound cue altogether. Without mutually-agreed standards, listeners must learn to navigate the task from scratch each time they encounter a sonification[11].

There is a recognised need for such cues to be specific to the task a listener is being asked to perform, and only to be used when necessary. Otherwise there is a danger that a

---

[11] Indeed, Smith and Walker (2002) caution that the improvement of task performance they observed with the addition of contextual cues might be explainable by the effect of practice.

listener might be overwhelmed. Even in relatively simple judgement tasks, the multitude of perceptual and cognitive subtasks a listener is required to perform to extract useful information from the sound stream is enormous. This passage from Smith and Walker (2002), speaking about share price sonification, illustrates the issue:

> "*When given the opening price and asked to report the price at a given time (noon, for example): the subject must listen to the entire graph, recall the pitch he or she perceived at approximately half the duration (noon time), compare it to the pitch perceived at the very onset of the graph (the opening of trading), estimate the change in price represented by the difference between the noon-time pitch relative to the opening pitch, and add or subtract that change*." (p. 4)

The task described here might be less difficult with the additional context of dynamic high/low points interspersed throughout playback, however these also represent additional sonic information to attend to. It is easy to see how attempts to facilitate the task in this way might actually increase the challenge of attending to the represented values themselves. The inclusion of a clicking *x* axis in this playback could lead to even greater difficulty, so if it is not absolutely necessary for the listener's task, it could be omitted.

It is clear from the approaches to auditory display reviewed so far that listening to sonified datasets can be a substantial perceptual and cognitive challenge. But this is only the beginning. Listeners may need to be explicitly instructed as to the nature of the contextual information provided in the task at hand and how they should use it. There is after all, no particular reason why a listener would know that certain tones are in fact reference tones representing dynamic high and low points. Furthermore, even if we assume that a given auditory display is structured clearly enough to enable effective perception of fluctuating sonic variables, and the task is within the attentional capacity of the listener, it does not necessarily follow that the listener will thereby gain any new knowledge or understanding of the subject of the dataset. The 'context' cues discussed in this section deal with the structure of a sound-listening task, not the broader context of the meaning-making task.

3.2.3 Auditory 'objects' and 'being' relative to them

Albert Bregman’s work on Auditory Scene Analysis (Bregman, 1990, 2001) is well-known in the field of Auditory Display and steps are usually taken to follow principles of sound design which will allow the listener to disambiguate auditory streams. However the application of Bregman’s work to auditory display may be misguided, as the ability to perceptually segregate sound streams to form representations of separate ‘objects’ is not in itself sufficient to afford understanding of a dataset - although it is sometimes treated as such (Saue, 2000; Scaletti & Craig, 1991). Sound streams in many auditory displays showcased at *ICAD* are conceptualised as meaningful objects in and of themselves, their makeup considered mostly from the perspective of psychoacoustics. The primary questions asked of mappings are those which will ensure that changes in the data stream will be reflected by perceivable changes in the sound stream(s) (Kramer et al., 1999). Indeed, sonic mapping decisions are seldom considered in the broader context of the task a listener is being asked to perform, i.e. to infer something about the world and perhaps make decisions for action (Fitch & Kramer, 1994; Flowers, 2005).

Walker and Kramer (2004) term this "conception and meaning-making" (p. 9) - among the three main subtasks for a listener which they identify (the other two being 'perception' and 'stream analysis'). Meaning-making has been sparsely accounted for in the history of auditory display research, especially in comparison to the other two identified subtasks.

Mapping to the sonic parameter of pitch is by far the most common strategy for auditory display (for a review, see: Dubus & Bresin, 2013). Perception of pitch change in a pure tone is extensively studied and its use has been validated in hundreds of sonifications, however in most cases it can be safely said that there is no meaningful link between a fluctuating pure tone and the underlying content/subject of the dataset. There is no culturally or ecologically-established reason why stock data would show up in the medium of a pure tone. This contingency is established by the sonification designer (perhaps in line with a discipline-internal standard) but may have been entirely unknown to the listener. For such a listener, a pure tone is almost completely devoid of behavioural relevance, except in some modern cultural-technological contexts, e.g. in societies which use handheld metal-detectors or car-reversal proximity sensors[12]. Perceiving anything useful about the stock market from this coded auditory information requires a significant cipher: propositional knowledge about the mapping - and the ability to put it to use in an attentionally-demanding interpretive task. Note that this example is a very simple mapping, computationally-speaking (one variable, one non-physically-complex sound stream), but this does not mean that gaining knowledge by listening to it is an easy task.

Going beyond the pure tone, whether amplitude, stereo-panning or any other auditory variable chosen arbitrarily with respect to the task (or for detectability first), parameter mapping can implicitly divorce real-world meaning from sound. The rationalist/objectivist focus on tightly-controlled stimulus design in psychoacoustics research is frequently imported to auditory display, arguably to the detriment of 'meaningful' auditory experiences. This focus brings with it a reduced conceptualisation of sound perception - as how controlled variation in objective parameters (frequency, amplitude, etc.) is tracked by the human auditory system generally. Sonification research in this vein tends to leave aside 'subjective' and fuzzily-defined phenomena like listener experience, the socio-cultural usage of sounds, moral and value judgements, politics and motivation - all of which are potentially rich sources of meaning for a listener (Supper, 2014). Worrall (2010) argues that this is characteristic of a kind of phenomenological dualism common in computer science (see also: Dourish, 2001), which assumes that mental representations of synthetic sounds are ontologically distinct from their use for an agent in the world. He further speculates that this makes some sense of the commonly-reported difficulty of achieving comprehension of even simple sonified datasets with few variables in ordinary (i.e. untrained) listeners. Often, parameter-mapped sonifications do not cater to everyday experience and world-involving skill, only listening ability and higher intellect. There is now growing interest in an

[12] And even in these situations, comprehension of the auditory information provided is inextricable from its use in the dynamics of a specific perception-action task.

'embodied' approach to sonification design to make auditory displays more generally understandable and widely adopted. The traditional approach of parameter mapping with a reliance on intellect and finely-tuned listening may be more suited to the niche group of experts who design sound than the general public.

3.2.4 Contextualising the meaning-making task

Recently, efforts are being made in some quarters to bring an up-to-date understanding of aesthetics into auditory display. The aim is that sonifications can be produced which are intuitively understood by untrained listeners, who can make use of their existing listening skills (Barrass, 2012; Roddy & Furlong, 2015a; Walker & Kramer, 2004). As Barrass (2012) puts it:

> "*... the aesthetic turn... [in sonification] moves the definition from 'interpretation' and 'communication' to 'usefulness' and 'enjoyment', reconfiguring it from an instrument of scientific enquiry to a popular mass medium for a broad audience. This approach also moves from engineering theories of information transmission to theories of the social construction of meaning*." (p. 5)

However, some ambiguity yet exists between aesthetics as 'meaning', and aesthetics as 'cosmetics', of which the latter is the more common usage of the term in auditory display. It is well-known that a pure tone display is difficult to listen to for long periods of time (e.g. see reports from participants in Schaffert & Gehret, 2013). Many sonifications designed with 'aesthetics' in mind are designed so purely to overcome this difficulty - to make sonifications pleasant to listen to. On the other hand, the recent turn towards aesthetic sonification in some corners of the auditory display community (as described by Barrass) draws from the pragmatist aesthetics of John Dewey (1934) and more recently, Johnson (2007), who argue that knowledge (or alternatively, meaning) is an inherently emotional quality, felt through embodied engagement with the physical environment and socio-cultural space. The proponents of this approach in auditory display are less concerned with low-level dimensions of sound, and more interested in how sound fits into our already-skilful repertoire of ways to engage with a personally-meaningful world.

Researchers in the auditory display community have long recognised that listener skill, culture, setting, mood and expectations can have a dramatic effect on how sonifications are interpreted (Ballass, 1994). It is only rarely that sonifications have been explicitly designed *for* these factors, rather than designed to sidestep them in pursuit of veridical communication. Misenheimer and Landreth (1993) designed their atmosphere and weather data sonification (named *Caustic Sky*) to draw on listener emotion directly in order to create a lasting impression regarding the poor state of the local atmosphere. Rain sounds indicated rainfall, while sirens and wailing sounds signalled sulphate concentrations and acid rainfall respectively. McCabe and Rangwalla (1994) sonified data from a fluid simulation using sounds from physically-modelled and recorded (real-world) water turbines. Certain types of sound can help to contextualise listening (i.e. aid education of attention) if the listener is experienced in their use for real-world tasks (Barrass & Kramer, 1999). Listening to the sound of water turbines is likely a much more meaningful experience for scientists working in fluid dynamics (the targets of that particular sonification) than a layperson. An expert should be much better-placed to detect subtle fluctuations in the sound based on its real-world relevance. If the intended audience of an auditory display is known, domain-specific options for sound use in sonification become available.

Roddy and Furlong (2014, 2015a, 2015b) have presented the most fleshed-out version of this approach to sonification design by leaning more heavily on its philosophical underpinnings. They argue that the 'mapping problem' in auditory display is really a manifestation of the 'mind-body/symbol-grounding problem' which has troubled philosophers of mind since Descartes:

> "*The misinterpretation of auditory cognition as the computation of context-free symbols transmitted along individual auditory dimensions is reflective of computationalism. This has led to the mapping problem... [which] requires a shift in focus towards embodied*

> *meaning-making, or, more accurately, embodied symbol grounding*." (Roddy & Furlong, 2014, p.5)

They argue that it is a mistake to think of aesthetics as - in any way - independent of a human observer armed with a particular repertoire of skills related to their particular way of engaging with the world. Such an approach could lead to the creation auditory displays which could only be sensed and processed by a computer, leaving real understanding unlikely (e.g. see: Searle, 1980). Related arguments pertaining to meaning come from Johnson (2007, 2015), who maintains that sensorimotor experience is the basis of pre-linguistic meaning and also more rational, higher linguistic reasoning and decision-making. Roddy and Furlong apply this school of thought directly to auditory display in the hope that it might inspire the creation of auditory displays which are understandable by all similarly-embodied listeners. It is useful to stress, however, that although most listeners likely share large similarities in their meaning-making faculties, no two experiences will be identical, nor should designers even aim for such.

> "*This pragmatist aesthetics perspective reconciles us to the assertion that user experience may only be designed* ***for****, that we must do all we can to maximize the opportunities for meaningful dialogue with our sonifications, but recognizing that the experience will not be universal*." (Barrass & Vickers, 2011, p. 160, emphasis mine)

While this recent aesthetic turn in auditory display has the potential to produce more meaningful engagements with sonification, it is difficult to formalise the approach into specific instructions for design. Barrass and Vickers echo the recommendation often-heard in design literature, that design should be iterative and responsive to the outcomes of user testing. Another difficulty for the designer is that rigorous empirical work to directly validate this approach has not been done - although perhaps it need not be. Roddy and Furlong (2014), in line with Barrass (2012), propose integrating this approach into sonification design according to the extent to which it can serve functionality and use of sonification in specific, situated task contexts.

3.2.5 Lessons for motor skill learning from Auditory Display

It was noted previously that theoretical discussions of sonification in motor skill learning and Psychology (where they exist) do not dwell on aesthetics. Like the equivalent treatments of aesthetics in Auditory Display, listener acceptance of sound is the primary concern (i.e. 'aesthetics' as cosmetics rather than meaning-making). Recent theoretical discussions around aesthetics in Auditory Display propose a fairly radical reorientation of the field and reinterpretation of what sonification should be - this reorientation moves the focus away from psychoacoustics and veridical communication towards real-world use of sound and emergent meaning. What the 'aesthetic turn' in Auditory Display demonstrates is that artificial sound does not need to be the sole 'carrier' of meaning; sonifications need not come prepackaged and self-contextualising, to be received and parsed by the listener. Aesthetically-oriented sonifications can guide education of attention and draw on practiced 'modes of being-in-the-world', wherein certain sonic patterns stand out as already-meaningful (similar arguments about scaling up the definition of a 'task' were made in section 2.5.4). This approach implicitly encourages a listener to *actively* engage with artificial soundscapes. In this way, understanding can be as much a product of the listener's experience as an agent in the world as it is due to the sonified content and design of the sound itself.

For motor skill learning, this approach opens up opportunities for the use of sound morphologies which are relevant in specific contexts, or in the task at hand. Depending on the task being trained, the context of learning and the learner's prior experience, certain ecologically-relevant sound types could be mapped to bodily movement instead of the more typical filtered noise and pure tones. Musical movement sonification, for example, could allow for this - as most people learn from an early age how a musical performance scenario works. Participants could thereby bring their own latent musical skill to an encounter with sonification. Understanding and making use of the information provided by sonification of movement could become an easier task, leading to reduced familiarisation times and faster improvement in perceptual-motor performance.

3.3 Sonification in Music and Interaction Design

The availability of electronic power and computation have provided the tools for the creation of new kinds of musicianship (Magnusson & Hurtado de Mendieta, 2007). Where before, musical instruments produced sound by way of physical excitation of some resonant structure, 'digital' instruments can now be programmed to produce sound in entirely novel ways. The defining feature of such instruments is that the causal chain of events which produces sound no longer needs to be entirely mechanical. Digital musical instruments are produced by cross-disciplinary teams of engineers, computer scientists, sonic artists and psychologists, who, unbounded by the constraints of form imposed by traditional instruments, aim to design the next generation of musical performances. The fruits of this research practice are showcased at international conferences such as *NIME* (*New Interfaces for Musical Expression*) and *ICLI* (*International Conference on Live Interfaces*). Given that the variety in form of digital instruments is so great (including even some installations which have barely any form to speak of), the rest of this chapter will selectively use the term ‘Digital Musical Interactions’ (DMIs[13]), as suggested by Gurevich and Fyans (2011), to incorporate the full spectrum of musical systems discussed in the literature.

Breaking and modulating the traditional relationship between the performer, instrument form and sound production has forced practitioners to confront the nature and mechanisms of digital interaction explicitly and directly. There is substantial theoretical and historical overlap between the concerns of contemporary DMI designers and those of designers more broadly interested in human-computer interaction (HCI). Indeed, *NIME* was

[13] This term was devised in order to provide a broader conceptual umbrella than that of more established terminology (such as 'Digital Musical Instruments/Interfaces') which orients discussion towards the physical device or its software. 'DMI' as used in this chapter incorporates the relations between performer, device and situation, and should not be confused with its other uses. Where this section engages in more narrow discussion of software or physical artefacts in and of themselves, more specific terminology will be used.

born in 2001 as a workshop at *CHI* (*Conference on Human Factors in Computing Systems*), and the fields remain closely aligned (Cook, 2001; Jensenius & Lyons, 2017). Both draw to some extent upon similar theoretical positions and ways of understanding the use of technological systems. What follows is a necessarily brief overview of current and historical discussions in the field of Interaction Design as they pertain to mapping, which is to a large extent inspired by the thesis of Nick Ward (2013), written on a similar topic. It is intended that the ideas developed here will inform the design of sonification mappings for human movement such that relations between interface, movement, sound, and situation can facilitate the development of skill. This section makes explicit the fact that moving with sonification is an interaction with a potentially alien technological system. The choices made by designers can facilitate or frustrate this interaction in equal measure.

3.3.1 Intelligence in interaction

The power to reshape the relationship between performer action and sonic output in DMIs has resulted in substantial academic interest in the relationship itself; the mapping. Instead of making sound by transducing energy from physical exertion into vibration of a resonant object, designers employ a wide variety of sensors and input devices to digitally capture performer movement for processing and use in sound synthesis. The mapping ‘problem’, i.e. the possibly infinite variety of solutions for tying movement and sound together - is not so much of a *problem* in Music, rather it is treated as a manifold of opportunities for creative interaction. If we treat the concept of ‘mapping’ as inclusive of: the form of the instrument/interaction, the sensors utilised and programming which transforms captured data into sound, then it could be said that ‘mapping’ is the central focus of many academic papers that focus on the design of digital musical interactions.

Suchman (1987) sets up a dichotomy between different ways of thinking about intelligent behaviour and associated knowledge (as they apply to computerised 'intelligent' systems) which serves as a useful primer for the forthcoming section. On the one hand is the kind of knowledge that can be written down, formalised for general understanding and used to construct an a-priori plan for 'purposeful' action. Suchman associates this brand of knowledge with the post-Enlightenment, Western ideal of knowledge as a rational construct which is context-free, can be transmitted from person to person, or traded as a commodity. To represent the use of this knowledge, Suchman employs the example of a European navigator who charts the ship's course ahead of the journey. Once en route, the navigator's job is a fixed, step-by-step procedure of checking the current position and prescribing course-corrections or changing the planned course. On the other hand is a kind of knowledge which cannot be expressed a-priori, but rather *emerges* in the context of the situation in response to its demands (see also the conceptualisation of embodied 'skill' developed in Chapter 2). Suchman's example for this case is of a Trukese sailor who has a destination in mind, but no course charted. Instead the sailor draws on whatever resources are available at the time (stars, sun, wind, tides etc.) to navigate and deal with problems if/when they arise. The sailor has a singular, only vaguely-stated goal of working towards the destination. These two kinds of intelligent behaviour are represented in the design of interactive systems, in that systems can require their users to implement either kind of knowledge.

3.3.2 Form and function

It has long been recognised in the field of interface design that the physical configuration of tools and interactive systems can elicit particular kinds of behaviour from the user. Hornecker (2005) states that:

> “*We can interpret [technological] systems as spaces or structures to act and move in, thereby determining usage options and behaviour patterns*.” (p. 4)

An encounter with an instrument or interface may be facilitated by whether the device looks like it can be bent, struck, or pressed, for example. More knowledge about the action possibilities afforded by a device emerges through use, by which the consequences of a certain kind of interaction can be evaluated. Sensory feedback, particularly tactile and auditory feedback, is fundamental to the development of skill in technological interaction

(Overbeeke, Djajadiningrat, Hummels, & Wensveen, 2000). A look back at the evolution of technological interfaces during the twentieth century can be illustrative of this point.

As technology and computing power have developed over the latter half of the twentieth century, there has been an identifiable shift in interface design and thus how we interact with technology. Øritsland and Buur (2000) provide a detailed illustration of this historical shift within the product range of a large manufacturing company. Over time, they note a tendency away from systems in which direct physical manipulation of mechanical controls was the norm to those which employ electronic, analogue controls, and finally, to systems which use virtual, on-screen controls. Ward (2013) argues that these three epochs of interaction style are representative of the wider electronics market (at both specialist and consumer levels) and are also identifiable in DMIs. The style of interaction for which devices are designed can have implications for skill acquisition.

3.3.3 Embodied interaction styles

Interfaces in the early-mid 20th Century were mechanical, consisting of "a single interface element and few operations" (Øritsland & Buur, 2000, p. 32). In this early mechanical era, there was a clear relationship between the movement of the user, the form of the device and its functional output; as the user practiced direct, bodily control over output. Øritsland and Buur explain that devices of this period required "intimate knowledge" (p. 32) to operate effectively, as the user's body was an integral and essential part of the functional system as a whole. Functions and their quality were related to the (sometimes heavy) muscular activation forces involved in the interaction. Many forms of sensory feedback were crucial in the development of this expertise, particularly the tactile sense. In the context of the task, the user and the machine were inseparable (Jensen, Buur, & Djajadiningrat, 2005). As an example, Jensen et al. describe the expertise and practiced skill of a worker in a beer factory, who works in concert with the moving systems around him to perform one of his tasks: manually checking to make sure an automatic failsafe system is operating correctly. They liken his practiced movements around his workspace to a dance,

"…*a 20 second choreography of precision and skill, one action following the other in one rhythmic, flowing movement*." (Jensen et al., 2005, p. 10)

This kind of interaction and the "intimate knowledge" required to operate devices successfully, can be taken as an example of Suchman's (1987) conceptualisation of 'emergent' intelligent behaviour. Successful use can be seen as enacted by the user, who is responsive to the conditions and constraints of the task in a moment-to-moment fashion. As Ingold (2001) would put it, it is the situated, intentional activity of 'making/doing' which itself discloses the function rather than reference to a fixed plan. Indeed, there is no plan which precedes the activity, other than some standard of what good output should look like - just like the Trukese sailor who can express his destination but not how he will get there. This is possible because of the continuous nature of the coupling between action and perception in the task. To perform the task effectively, the user must learn dexterity, care and judgement in response to the ever-changing conditions of the task (Ingold, 2001, p. 21). It is important to note that learning the 'intimate knowledge' required for this task (i.e. acquisition of skill) requires extended physical practice.

A focus on the body and 'skill' are common themes in the ever-growing literature on embodiment in interaction design which, broadly speaking, pushes for a greater role of the body and its impressive capacity for skilful action in the design of devices (Dourish, 2001). There is an aesthetic component partly driving this movement; Djajadiningrat, Matthews and Stienstra (2007) argue that interactions which are functionally dependent on bodily movement can be pleasurable, in the same way that playing an instrument or creating physical art can be pleasurable. They also maintain that this 'beauty in use' is not dependent on devices being *easy* to pick up and use, in fact sometimes the opposite is true. No-one ever learned to play a violin because it was easy, but the achievement of skill or mastery makes playing an undeniably pleasurable experience. For design, Djajadiningrat et al. propose that digital devices may better cater to perceptual-motor skill by instantiating functional and use-relevant information in physical form, to which the user can learn to become sensitive. In the earlier beer factory example, the layout of the factory floor, the shapes of the multiple tools and the movement of the conveyer system all support the development of bodily skill and dexterity, purely by virtue of being a space in which it can develop.

It is informative to contrast this style of “embodied interaction” (Djajadiningrat et al., 2007; Dourish, 2001) with the multifunctional but movement-constrained interactive spaces offered by modern systems which have seen widespread adoption. Much of modern interaction design has become detached from the capacities of the body, relying more and more instead on those of a 'higher' mind. The contrast with embodied interaction can highlight some of the central issues faced by sonification designers, who hope to marry the potential of modern computation with bodily skill in the creation of DMIs.

3.3.4 Intellectual interaction styles

A rise in computation and electrically-powered components in the late twentieth century allowed many tasks to be performed automatically, and the user became a director, acting through what had become a much more 'general' interface (D. Norman, 2002) - one with a single interface and many operations. Today, human interactions with technology - now mostly computers of one form or another - have taken the form of button presses and manipulation of on-screen menus, regardless of the function to be performed.

Modern devices are built to be multifunctional and so much of design caters to ease of use above all else (Overbeeke et al., 2000). A learning curve in technological interaction is an obstacle to be overcome by designing interfaces simple enough that anyone can use them immediately (a testament to the success of this philosophy is the growing adoption of iPads by babies and very young children, see: Hourcade, Mascher, Wu and Pantoja, 2015). The most ubiquitous example is the modern personal computer[14]. Its functions are wide-ranging, from simple tasks like word-processing and searching for information, to the more complex, e.g. video editing and music production. Despite the potentially infinite variety of possible functions, our physical interactions with computers are limited (in most cases) to those

[14] A computer can also be considered a digital instrument (Cascone, 2004). Most often, computers only form an intermediary between sensor input and sonic output, but in some cases, such as ‘live coding’ performances, the computer can be considered the whole of the instrument.

afforded by mice, keyboards and more recently, touchscreens. Tactile feedback is generic and does not vary based on effective completion of the task. Visual feedback is specific, but is mostly geared towards intellect (e.g. messages), telling the user symbolically, through language or learned metaphor, whether a function has executed successfully (Bellotti et al., 2002). The crucial change which characterises the modern era of interaction, however, is to a reliance on the cognitive system. Users now perform functions by following remembered instructions and implementing step-by-step plans. Intellect is required rather than bodily skill (Ward, 2013).

The central issue for interaction design is best summed up by Overbeeke et al. (2000):

> "*What happens inside electronic products is intangible: it neither fits the mechanics of our body nor the mechanical view of the world. In contrast with mechanical components, electronic components do not impose specific forms or interactions for a design. Products have become 'intelligent', and intelligence has no form*." (p. 1)

Without form, the possibilities for physical engagement with devices are diminished (Hornecker, 2005). Indeed, most modern software interaction requires some use of Suchman's (1987) first kind of formalisable 'intelligent' behaviour. Performing a task requires selecting a functional goal ahead of time, recalling a list of steps (e.g. navigation through menu trees or commands) and formulating a plan to be serially enacted by the motor system. When a mistake is made in the task, it is not typically due to clumsy bodily movement - as the movements required are designed to be so simple that a baby can implement them - but due to an error in the plan. Responding to unexpected circumstances is a protracted and physically detached affair requiring the interpretation of symbolic feedback and reformulation of the plan - which can then be passed to the body and implemented anew. There is an implicit mind-body split in modern interaction with computerised systems: an assumption that a rational, disembodied mind is the controller of the body (Dourish, 2001). The body is considered in modern software inasmuch as it can be used as a tool to carry out the commands of the mind, providing inputs for the device. Donald Norman (2002) convincingly argues that this trend in design has led to the ubiquity of devices which are difficult to use because they place an excessive burden on the implicitly disembodied mind. His most striking example is that of an office telephone system which has a wide range of functions, but only the standard 12 buttons, plus a couple more. Functions such as putting a caller on hold, merging calls etc. are possible in principle by inputting arbitrary 3-digit codes - but no-one in the office can remember them, so these functions are never used.

The arguments against this style of interaction are more than cosmetic. In contrast with the 'embodied' interaction style explored in section 3.3.3, what I have taken to calling the 'intellectual style' requires that most of the functionality happens away from the body. Compared to the responsive and mechanically-structured work environment of the beer factory, these are not interactive systems in which bodily skill is integral to use. Learning to use the system is fundamentally a cognitive challenge, and competence lasts about as long as the user's ability to remember an abstract plan or series of instructions.

3.3.5 Mapping movement to music

Technological interfaces afford the development of perceptual-motor skill when the user is placed in direct, tactual-perceptual control of the unfolding function. Ward (2013) applies this position to an analysis of DMIs, and finds many modern implementations lacking in opportunities for the development of skill. With the proliferation of cheap computing power has come the standardisation of interface elements (buttons, switches and sliders) and displays with which a user might synthesise sound. The prototypical (and most widely-used) example of a device for the creation of digital music is the personal computer, which limits physical interaction possibilities. The poverty of expressive movement and engagement of the senses in DMIs is greatly lamented in some circles, as typified by Norman, Ryan and Waisvisz (1998):

> "…*how we evolve in the digital-physical world essentially depends on our dealings with keyboards, mice, joysticks, and touchscreens... screen athletes are praised in all their ponderous immobility, and the delights of real movement are insidiously overridden*."

3.3.6 Sonic constraints on (inter)action

This section has to a large extent dealt with physical constraints on behaviour. That is, how the physical layout of an interface solicits certain kinds of actions, and how the relationship between action and function can constrain interaction strategies. However to fully appreciate digital *music* production as an interaction, it should be recognised that performers are coordinating with sound, which itself has a causal role in behaviour. In an explicitly Ecological analysis of musical performance, Windsor and de Bezenac (2012) propose that the concept of 'affordance' can be a useful way to think about behaviour in a musical context. Affordances - although examples of the concept are often described in visual terms - are not modal; the opportunity for a certain kind of action can be perceived using auditory, visual or tactile sensory information - or any combination thereof (see section 2.3.1). Digital musical instruments afford the production of sound. The various kinds of sounds which can be produced effectively narrow the range of actions which might be performed. To understand this, it helps to remember that according to the Ecological approach, action does not precede perception (nor vice versa). Actions are not planned in isolation from the perceptual information they will produce, rather, each implies the other. Producing even a simple melody or rhythm is therefore an emergently purposeful task[15], as an infinite array of possible behaviours stretching into the future can be collapsed down to those which will continue or complete the musical pattern under the current circumstances (Todorov & Jordan, 2002). Music-making behaviours, however they are enacted (using sliders, touchscreens, reeds or sticks) are differentially potentiated by the context of the ongoing performance.

Situations also in part specify appropriate music-making behaviours, depending on the social or coordinative *purpose* of music. Krueger (2014) provides many examples of music as a tool for social coordination and emotional modulation - for liveliness at parties, the opposite at funerals, focus for studying, elevating arousal for athletics and many more. As

[15]As opposed to a-priori purposeful.

such, some forms of music-making are potentiated by situations, and others are out of the question. This grain of analysis reveals that performers do not simply act to produce sound for its own sake. Indeed, taking a wider view of musical interaction as structured by the social and cultural context can in fact help explain the form taken by the lower-level sensorimotor activity we have heretofore been interested in.

However, the kinds of sounds available, how one might act to elicit/modulate sound, and what kinds of sound are desirable in the first place may not be immediately obvious upon an encounter with a new, computerised musical instrument (as compared to an encounter with a familiar acoustic instrument). Ecological theory predicts that substantial experience is required in order to learn to detect the many subtle affordances available in a musical interaction (E. J. Gibson & Walker, 1984; Krueger, 2014). An interesting feature of digital musical interaction, where it involves novel devices, is that there might not be a strong culture of musical practice available which can be drawn upon to help constrain music-making behaviour. Affordances therefore, might effectively be 'invisible' to begin with.

3.3.7 Understanding musical interaction

Theorists interested in the design of new musical instruments and sonic interactions have elaborated several ontologies to better understand how performers interact with musical-technological systems.

Chadabe (2002) suggests that all digital instruments exist somewhere on a continuum between total determinism of sonic output and total unpredictability of output. Deterministic instruments are those in which no new information is added by the computational element of the system; the performer is in total control of sonic output via their movements, much as with traditional acoustic instruments. At the unpredictable end of the spectrum, a large amount of control over sonic output is given over to algorithms which elaborate on inputs in unpredictable ways and systems with internal feedback loops which produce unplanned, emergent sonic behaviour (e.g. Di Scipio, 2003).

Gurevich and Fyans (2011) report a study showing that with the use of a more deterministic mapping, the performer's bodily skill is generally more understandable to naive audiences. Cadoz (2009) argues that an informative perceptual 'symmetry' between performer action and sonic output is possible in digital instruments by respecting (roughly speaking) the transfer of energy between real-world mechanical systems; in other words, mappings can make sense to performers and audience members not versed in the programming of the device by imitating (or simulating outright, through physical modelling) the dynamics of real-world action-sound relationships. So-called 'instrumental' interactions are those which are more suited to the traditional acoustic performance scenario, where the intent of the performer can be easily interpretable by the audience (Fyans, Gurevich, & Stapleton, 2010; Rodger, Craig, & O'Modhrain, 2012). Less-instrumental interactions can render the control mapping effectively invisible, leaving naive audiences unable to interpret the performer's intent from their movement. As Cascone (2004) rather cynically puts it,

> "…*the standard codes of musical performance are violated: the laptop is doing the work, no skill is required or demonstrated, and the artist could just as easily be any one of the audience faking a performance*." (p. 103)[16]

However, it should not be understood from this analysis that 'instrumental', or skill-dependent interactions can only be those with a straightforward, 'symmetrical' mapping between input and output. Digital musical interactions, like acoustic instruments, do not necessarily have to be easy to control. In fact, some resistance to control may be desirable. An informal experiment reported by Hunt, Wanderley and Paradis (2003) reveals that mappings with some complexity can encourage exploration and the development of fine-grained skill (as compared to simple direct mappings, which are 'figured out' relatively quickly). The authors describe two versions of a simple instrument, which used exactly the same interface and sound type (two sliders which control the output of an oscillator). In the

[16] In the use of the terminology "standard codes", Cascone implicitly acknowledges that judgements of musical 'skill' or 'acceptability' are entirely specific to the musical practices recognised by the perceiver.

first version, one slider controlled volume, the other pitch. When using this version, participants/performers were only interested in playing for around two minutes, as the mapping was simple enough that its use quickly became trivial. In the second version, one slider needed to be in continuous motion for the oscillator to produce sound at all, and its acceleration controlled amplitude. The same slider also controlled pitch, but in the opposite direction to what participants expected (sliding down raised pitch and vice versa). The second slider also modulated pitch, but more subtly than before. Using this more complex mapping, participants were found to become much more engaged in working with the device - and motivated to develop the skill to control sonic output. Most persevered with this version of the instrument for longer than version 1, and nearly all reported the experience as "rewarding" and "more like an instrument" (p. 2).

Wessel and Wright (2002) address the same issue of difficulty and skill in mapping design with the use of a floor/ceiling metaphor. They see a low skill 'floor' as desirable in musical interaction, i.e. it should not be difficult for a novice performer to 'get started' with playing. However they highlight the same issue as Hunt et al. (2003): that often instruments designed to be accessible to novices also have a low skill 'ceiling' (like the first slider device described above):

> "*...many simple-to-use computer interfaces proposed for musical control seem - after even a brief period of use - to have a toy-like character. By this we mean that one quickly "out-grows" the interface by discovering the limits of how it can be used. Many such simple-to-use interfaces do not invite continued musical evolution*." (Wessel & Wright, 2002)

More complex mappings can raise the skill ceiling of an interaction - allowing even for the development of virtuosity - but may make the instrument inaccessible for novices. The authors propose that a way around this problem could be to employ physical controllers where possible, which invite actions metaphorically aligned[17] with the sonic output of the system. This strategy could maintain a high skill ceiling while keeping understanding of use grounded in familiar physical interactions.

Other theorists have elaborated taxonomies to describe the ways in which *spectators* understand musical interaction with new technology. Reeves, Benford, O'Malley and Fraser (2005) propose that the spectator experience can be described by the interaction's position on two dimensions (see Figure 3.1). These dimensions characterise 1) the perceived energy in the manipulation enacted by the performer, from low/hidden (pressing a button) to high/amplified (large, full-bodied movements/dancing) and 2) the perceived strength of the output, from hidden (on-screen messages seen by the performer) to amplified (loud sounds, or large movements of noisy prostheses, in the case of Stelarc).

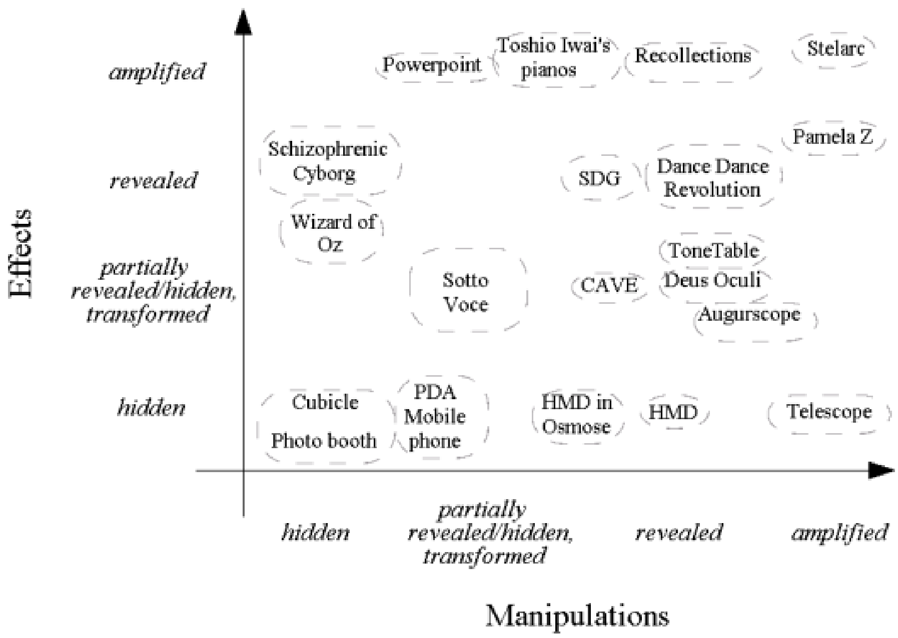

Figure 3.1: A taxonomy of interactions, as perceived by spectators. On the x-axis is the required functional movement of the performer (manipulations), and on the y-axis is the perceived effect of that movement (Reeves et al., 2005).

---

[17] This suggestion draws upon embodied conceptual metaphor theory attributed to Lakoff and Johnson (1999). For Psychological context, see section 2.7.5.

Where on this taxonomy a given interaction falls can reflect the convergence of the performer/designer's expressive purposes and the affordances of the technology available.

The ontologies of interaction discussed so far provide clear ways of thinking about the performer's experience: how they think about their instrument and what they can get out of it. Equally, they can help make some sense of how spectators or audience members experience the system. What most work discussed so far has left relatively implied, are the social and cultural environment within which musical interaction occurs. In line with the Ecological analysis of Krueger (2014) and Windsor & de Bezenac (2012), Waters (2007) argues that a musical performance can only be fully understood as an non-decomposable ecosystem, constituted in relations between performer, instrument and environment. According to this view, which stems from critical social theory, 'agency' (or, intentionality) in musical performance does not come solely from the performer, but can fruitfully be seen as embodied in his/her interaction with a widely-networked and distributed system (Clark & Chalmers, 1998; Gallagher, 2013). This is one of the most inclusive ontologies of digital musical interaction - the system here described can incorporate the performer's movement, available materials (including their affordances and resistances), the sonic mappings, socio-cultural context and any available virtual or physical spaces. It is from the interaction of these interconnected resources that a performance emerges - and within which an inclusive analysis of skill can be conducted (Ingold, 1996). An aim of the ‘ecosystemic’ approach to musical interaction is that the explicitly distributed nature of agency in the musical systems designed (e.g. Stapleton & Davis' *Ambiguous Devices*, which is a networked, partially autonomous musical system distributed across two locations; see also examples in Waters, 2007) can structure thinking about performance generally as a spatially and socially distributed system (Magnusson, 2009). Gurevich, Stapleton and Marquez-Borbon (2010) designed a musical device at the other end of the spectrum of device complexity for a qualitative study, which also demonstrates the approach – a simple box with a single button on it which activated a pure-tone oscillator, locked to one frequency. The researchers asked a group of music students to spend an extended time learning to use the instrument and to come up with a performance. Despite the highly constrained nature of the device (and, the authors argue, *because* of it - to some extent), large variations in style emerged during performance, as performers drew upon a variety of musical skills and experiences to shape the interaction (see Figure 3.2 for examples). It is proposed that the same interaction philosophy can help to make sense of the enormous possible variety and idiosyncrasy in musical performance, i.e. the form of the interaction is *not only* constrained by the form of the device and how it maps movement to sound, but by the distributed sociocultural resources drawn upon by the performer him/herself.

Posture	*Ways of holding*	*Ways of playing*	*Musical variations*
Sitting down	Box on table	Button press with finger	Rhythmic beeping
Straight back	Box on lap	Button press with thumb	Arrhythmic beeping
Leaning forward	One-handed	Finger tap on box	Rhythmic tapping
Arms to side	Two-handed	Thumb tap on box	Arhythmic tapping
Arms projected	Held by length	Hand tap on box	Sound filtering
Elbows free	Held by width	Manual filtering	Mechanical noise
Elbows resting on lap	Box rotated on any axis	Spatialization	Simultaneous events
Both feet flat on floor	-	Use of power switch	-
Legs crossed	-	Compound gestures	-
Standing up	-	-	-

Table 3.2: Interaction styles with a highly constrained instrument. From Gurevich, Stapleton and Marquez-Borbon (2010).

3.3.8 Interaction summary and lessons for mapping in motor skill learning

This section has covered the various ways in which interaction with technological and musical-technological systems can be viewed. To begin with, the dichotomy between so-called intellectual and embodied forms of behaviour was crucial to explain how the processes of digitisation and computation have altered our relation to technology, and to the same extent, music. An examination of the user's action and the degree to which they are in direct control of the output of a system reveals that to some extent, bodily skill and function have become separated by computerised technology (Ward, 2013). The modern user of music-technology can be forced to act more like a manager, who must plan abstractly, than a craftsman or traditional musician. It might be the latter form of interaction, constituted in continuous perceptual and bodily engagement with the material[18], which better affords both creative use and stable motor solutions to the problem of technological interaction (Todorov & Jordan, 2002; see also section 2.2).

However, the latter half of this section considered theoretical discussion in the context of technological musical interaction specifically. Here, we see that the nature of mapping between movement and sound, while critically important for novice and spectator understanding is not the whole picture. The idea of 'emergence' is further highlighted by critical social analysis, which constructs an inclusive definition of 'skill' in performance as a product of the socio-cultural context as much as the material context.

3.4 Chapter summary: An approach to motor skill learning with sonification

The immediately previous discussions on digital musical interaction highlight that: 1) a coherence between movement style and sound is critical to novice performer and audience understanding of digital (musical) interactions, and 2) The development of skill in interaction is facilitated when musical interaction is allowed to take place 'in the world', via continuous perception-action interplay within a broadly-defined culture of practice, rather than predominantly 'in the head', or predominantly inside a device. Both of these notions are informative for the development of an approach to mapping sonification for motor skill learning.

A learner must be able to understand how their actions are functionally related to the sonic output of the system, but this is not a guarantee where digital sound is involved. On the one hand, it is realistic to expect that some preliminary practice could be required for a learner to discover how their actions modulate sound. This could even work in favour of motor skill learning - if education of attention towards subtle nuances in sonic feedback is coupled to the control of task-relevant qualities of movement, then there is scope for

[18] 'Material' in this context can include sound, a case made by Steenson and Rodger (2015).

auditory and motor learning to work in parallel, each effectively augmenting the other. This is after all what happens during learning of an acoustic musical instrument. However, where the aim is to teach a *motor* task per se, perhaps sonification need be only as complex and nuanced as necessary in order to highlight the relevant aspects of movement which should be controlled in competent task performance. An entirely novel interaction with a new motor task and sonification feedback system could be facilitated with a straightforward, deterministic control mapping. There may even be a place for a tiered approach, in which more complex and difficult-to-control sonification mappings become available upon attainment of a more basic level of motor competence. Learning to control the more complex mapping could then scaffold development of more accomplished motor control, while also stimulating engagement (Hunt et al., 2003).

The affordance-based approach to digital musical interaction has some direct applicability to motor skill learning with sonification. Recognition of music-making as an activity which carries its own emergent intentionality lends some more credence to the pseudo-musical movement sonification strategy explored in section 3.1.5 - in which practice with ‘direct sonification’ is structured like a music lesson and the learner's job is to 'play' the task correctly, as demonstrated by an exemplar. This strategy helps to constrain movement in a moment-to-moment fashion. The next move for the learner is always known - it is the move which will continue/complete the ongoing musical pattern. This brings sonified motor skill learning close to something like the 'Suzuki method' of musical instrument training, in which children learn to play a piece primarily by ear, after becoming intimately familiar with how it should sound.

In an earlier section (3.2), I suggested that sound designers in Auditory Display are faced with a particularly difficult challenge in needing to convey knowledge through sound alone - whereas in Psychology, human movement is inherently meaningful to the mover and can help contextualise sonification to some extent. This idea of 'sound alone' as the main carrier of knowledge is fundamentally flawed, as it fails to recognise that knowledge must be grounded in relation to the world to be meaningful to an embodied agent (Johnson, 2015;

Roddy & Furlong, 2014). Relying on sonification to 'carry' a lot of information (possibly including internally-oriented[19] contextual cues for interpretation of the soundscape), is effectively front-loading cognition and intellectualising the task. This is an approach which is sometimes adopted by sonification designers in Psychology, where motor skill learning can be conceptualised as the elaboration of an 'internal model' or 'representation' (e.g. Effenberg et al., 2016; Oscari et al., 2012), and sonification can be thought of as a series of data packets transmitted to the rational controller of the body (take 'error sonification' for example). The 'Aesthetic Turn' in Auditory Display (Barrass, 2012) seeks to redress this historic dualism by bringing the listener and their embodied skills into the encounter with sonification. Ecological approaches to musical interaction similarly stress the distributed nature of interaction with sound, in which the performer has a rich and varied set of worldly resources to call upon. What Psychologists can take from this approach is a respect for the learner as an embodied agent, with skills beyond the ability to reason logically and perform calculations. With the use of sound morphologies which are meaningful in the context of the learner's everyday lived experience (as opposed to sine tones and white noise), it might be possible to guide his/her attention towards the auditory informational patterns which are maximally important - which if controlled, (assuming they are mapped to task-relevant aspects of motor performance) could boost motor performance.

A particularly interesting example of movement sonification aesthetics from the field of HCI which illustrates this point is reported by Stienstra et al. (2011). The authors mapped pressure on different parts of the skater's boot to pink noise, which was modulated in terms of amplitude, bandpass filtering and central frequency depending on pressure distribution. The result was that the skater could hear informationally-rich wind-like sounds when she skated. The authors describe this strategy as "non-coercive", in that there is no clear 'ideal' type of sound to aim for; the mapping is mainly meant to enhance information pickup from

[19] Internally-oriented as in cues which exist *in* a sonification, to guide interpretation *of* a sonification, like the dynamic high/low-point tones in Smith and Walker's (2002) share price sonification.

the tactile sense in the feet. The skater was allowed to discover the relationships between her own movement and changes in the sound. Granted, this took rather a long time (8 hours overall), but verbal reports from the skater were reported as a display of "outspoken enthusiasm" (p 42). After familiarisation, it was clear that the sonification was useful, as the skater was able to exert much finer control over her weight distribution when sound was available and was skating more efficiently. The authors note,

> "*To listeners who never used the ASE [augmented speed-skate experience], the sound-scape recordings of the sonified movement sound like meaningless wind. To Frouke Oonk [the skater], on the other hand, this wind caused by her movement provides a continuous rich information flow*." (p. 42)

This description reveals starkly the extent to which sound perception in sonification is intertwined with bodily skill and interactive experience. In this case, the abstract nature of the sound morphology used and the unfamiliarity of the system in general necessitated a long familiarisation period, but it is clear that after practice, the skater was able to bring the sound into use for greater perceptual-motor control in her task.

Going beyond the design of sound for a moment and considering interaction, it seems that an approach to sonification commensurate with the perception-action stance developed in Chapter 2 would be one which merges the notion of 'understanding sonification' with embodied 'use' of sonification. In line with Steenson and Rodger (2015), I would argue that the meaning of sound is inextricably a part of how it is (or can be) used in worldly tasks. Furthermore, motor control and the acquisition of skill with sonification can be characterised as the intentional 'bringing into use' of dynamic relations between body and physical/sonic/sociocultural environment (Ingold, 2001). The 'working out' of sonification as feedback can be thought of as happening largely through interaction with the world rather than solely in the mind.

3.4.1 Limitations of the approach

It is not as simple as to say that ‘embodied skill’ can *only* present itself in deterministic, instrumental sonic interaction where movement and sound are tightly coupled and the learner's own skill plays a role in active listening. In their analysis of a wider range of DMIs, Gurevich and Fyans (2011) make room for skill in both ostensibly embodied and intellectual interactions (like live coding performances), while recognising differences in the nature of each task. Similarly, Rodger (2010, see section 3.1.2 therein) convincingly dissolves the distinction between intellectual and perceptual-motor skills with reference to common underlying neural mechanisms and ontological flexibility. A performer can have a meaningful experience and exercise skill in almost any kind of sonic interaction. Indeed Dreyfus (1996), in his phenomenological analysis of chess-playing (which is inarguably an intellectual challenge to the novice), shows that through extensive experience, intellectual tasks can effectively become perceptual-motor tasks, as to-be-remembered rules and procedures become incorporated into the learner's practiced habits of perceiving and acting. In this way, the prior symbolic, 'rational' plan for action becomes grounded in enacted skill, to the point that the learner (now expert) may no longer be able to propositionally state what his strategies are.

The reason that I have largely filtered this chapter through (very loosely) Suchman's (1987) dichotomy between emergent and intellectual knowledge is that this same dichotomy is reflected in the field of Psychology. Embodied/Ecological and Cognitive approaches to Psychology, while they address the same topic, often ask very different questions and perform very different experiments (Clark, 1999; Wilson & Golonka, 2013). There can be an assumption that perception-action and cognitive processes, as ostensibly distinct psychological constructs, are the purview of the former and latter respectively (Chemero, 2009). In order that this chapter and its approach might steer the field (as a whole) towards more productive outcomes, I have tried to harness orthodox thinking.

At a more practical level, while perceptual-motor and cognitive tasks may not require truly distinct psychological processes, I have attempted to highlight as many ways as possible to ground the practice of sonification in embodied capacities for meaning-making in the hope that the approach might be more generalisable. The idea of a fully generalisable system of aesthetics is problematic, in that it is unrealistic and probably undesirable (Barrass & Vickers, 2011; Johnson, 2007), however it seems likely that an embodied style of movement sonification might be more widely acceptable for users than a symbolic, intellectually-challenging one. By the logic of the current approach, human beings share similarities in body shape and their capacity for interaction with the world; therefore knowledge which can emerge through interaction alone may be more accessible to more people. It is important to remember however that a given implementation of sonification can and will mean different things to people with different backgrounds, experience and skills (Roddy & Furlong, 2015b). If this fact can be harnessed on a case-by-case basis to create more meaningful interactions, all the better.

3.4.2 Chapter conclusion

If sonification with motor skill learning is designed as an interaction which exists in the world (i.e. is enacted by the body and perceptual system in concert with the interface and existing skills), mapping should be more intuitively understandable and afford the development of perceptual-motor skill. Relations between the interface, the movement of the user and the embodied resources he/she can make available to guide performance are vital. Sound production should not be seen (by a designer) as the end of a linear chain of events (planning > movement > sound, then a loop back to the planning stage), rather it should be seen as a loop in which movement and sound inform each other continuously (movement <> sound). Offline planning need not be a part of the interaction as designed; its form (both movement and sound produced) should emerge in the context of a perception-action task.

Chapter 4

Movement sonification as feedback for learning a new task: implications of misguidance

4.1 Abstract

Emerging evidence suggests that movements which are sonified (i.e. transformed into sound with imperceptible latency) can be perceived and controlled more effectively than the same movements performed in the absence of sound, with sonification acting as concurrent, augmented feedback. The current experiment aimed to test the application of sonification to the learning of a new motor skill. Participants (total N = 45) learned a table-top bimanual task which required the timed manipulation of three handheld objects. In one condition, movement trajectories and object arrivals at target zones were sonified. In another, only object arrivals were sonified. Performance and learning (delayed performance testing after the end of practice, with feedback withdrawn) between these and a third control condition (in which participants practiced the same task without sonification) were compared. Benefits of either kind of sonification were not found; sonification may even have hindered performance in the early practice stage. Additional analyses indicated that sonification may have placed unexpected constraints on motor performance which were not present in the control condition, leading to a mismatch between participant-perceived task goals and those measured. In an extended discussion section, the experimental task is re-examined in terms of constraints on perception and action in bimanual coordination tasks.

4.2 Introduction

4.2.1 Learning and feedback

Sonification of movement (the use of technology to turn movement into sound) has gained momentum in recent years as a tool for movement performance enhancement in sporting, laboratory and rehabilitative skill (re)learning contexts (Rosati et al., 2013; Schaffert, Mattes, Barrass, & Effenberg, 2009; Schmitz & Bock, 2014). This recent interest has the potential to place sonification on equal footing with more well-established forms of augmented feedback for motor skill learning, such as verbal instruction/feedback, graphical display of performance data and numerical scoring. The longstanding consensus on the effects of augmented feedback propounds that such extra information provided during practice can have enhancement effects on both quality and speed of skill learning (Magill, 2011). However numerous studies show that the provision of such feedback often leads to dependence on the ‘guidance’ it provides (Maslovat et al., 2009; J. H. Park et al., 2000; Sigrist et al., 2013a). Learners come to rely heavily on artificially-provided augmented information to judge the outcome of their motor performance, and struggle when it is removed (see section 2.6.3). This 'guidance effect' is obviously not ideal if training is to transfer outside of laboratory conditions, e.g. to the field, racetrack or everyday environment, where artificial feedback systems are either forbidden (as in sports) or unavailable for practical reasons.

More recently, it has been suggested that the mechanism for the guidance effect can be found in lack of processing of task-intrinsic sensory information in favour of information provided by augmented feedback (Sigrist et al., 2013a). When making judgements about movement events (including our own), the sensory information picked up from the environment is weighted according to its reliability (Ernst & Bülthoff, 2004). More reliable information is weighted more highly than less reliable or 'noisier' information. In general, integration of perceptual information from across multiple sensory modalities (vision, sound, haptics, proprioception etc.) leads to a synergy of information, and more accurate percepts of performance (De Gelder & Bertelson, 2003). However, given that augmented feedback information can be extremely accurate and reliable, it may be weighted more highly by the learner than task-intrinsic information, such as proprioception. The net result of these processes is that attention is captured by the feedback display during learning, leaving intrinsic information neglected. It should be a requirement of transferrable learning for practice to include experience of proprioception and other intrinsic information sources. Intrinsic sources after all, are all that a learner has access to following the removal of feedback. Demonstration of improved motor performance which is not dependent on the immediate presence of augmented feedback is a fundamental aim of the current experiment.

Traditional attempts to ameliorate the guidance effect of augmented feedback have involved training with different schedules of feedback. For example, feedback can be provided on some trials and not others, or more often in the early stages of learning than the later (D. I. Anderson, Magill, Sekiya, & Ryan, 2005). The rationale behind these approaches is to force the learner to gain task experience using only intrinsic information sources, while still availing as much as possible of the enhancement effects of augmented feedback. Under these arrangements, augmented feedback can help to guide performance when it is available, but learners do not have to adjust to an entirely unfamiliar situation when feedback is withdrawn following the practice stage.

Although the guidance effect is a well-established psychological phenomenon, appearing in Psychology textbooks on motor skill learning (Magill, 2011), the empirical research behind it is in fact rather old. Many of the experiments cited in evidence of the effect investigated only augmented feedback presented either verbally or graphically, and most authors of the time carried dated assumptions about wholesale internal representation of motor skills based in muscle activations (Adams, 1971; Chamberlin & Magill, 1992). Little attention was given to the nature of feedback information itself, and findings were assumed to apply across contexts and information modalities, despite the fact that most relevant research employed vision alone. Consideration of the structure and presentation method of visual information provided in classic augmented feedback experiments in fact produces some novel insights about feedback and the nature of the guidance effect.

4.2.2 Information presentation and structure

Augmented visual information of the kind used in classic motor skill learning experiments is not 'augmented' in a way which would be familiar to a modern user of augmented reality technology. That is to say that these experiments historically have not presented visual information overlaid dynamically on the visual scene. Most employ a separate screen, on which a graph representing performance (or deviation from correct performance) is drawn (Vander Linden et al., 1993). The learner is encouraged to monitor movement performance by visually attending to the display. In arrangements which present visual feedback concurrently with practice (i.e. the graph is generated 'live' from user movements), learner visual attention is plainly distracted from intrinsic information and performance-monitoring outside of what the feedback provides (Kovacs & Shea, 2011; Maslovat et al., 2009; Wang, Kennedy, Boyle, & Shea, 2013). That this would result in attentional capture by augmented information and neglect of intrinsic sources is no surprise. Dependence on feedback systems arises because learners have only learned to use feedback to monitor performance - neglecting the sights, sounds and feelings of performing the task without feedback. Alternate schedules of feedback delivery (feedback only on some trials, phasing out feedback, etc.) do make some sense as a strategy to deal with augmented feedback dependence and the guidance effect in this case, as they force the learner to gain experience with task-intrinsic sensory sources (J. H. Park et al., 2000; Winstein & Schmidt, 1991). An interesting feature of sonification as concurrent augmented feedback is that it can potentially stand in as a more effective alternative to altered scheduling of visual feedback. The human auditory system is suited to the pickup of information from the environment beyond the range of vision. Sonification as feedback does not require the learner to look away from the visual aspects of the basic motor task. Therefore learning a new task with sonification as feedback should allow learners to make use of augmented sonic *and* intrinsic visual information concurrently. Alternate scheduling of feedback may not be necessary - learners may be able to avail of the full benefits of sonification as feedback while concurrently monitoring task-intrinsic perceptual information, perhaps thereby alleviating the guidance effect.

While undoubtedly part of the story, attentional capture alone does not fully explain the nature of the guidance effect in visual augmented feedback. The potential task-altering effects of variable *information structure* are often overlooked in augmented feedback investigations, both historically and in much of the contemporary literature (for examples, see a review by Sigrist et al., 2013a). Information structure is conceptualised here as the perceived Gestalt form of information; it can also be said that structure defines how information can be used by learners. Visual augmented feedback, which is typically presented as a graph, is a structurally-transformed version of the intrinsic visual-kinematic information which inarguably must be picked up and used during practice for learning to transfer outside of feedback conditions (e.g. Vander Linden et al., 1993; see also: Wilson, Snapp-Childs, Coats, & Bingham, 2010).

In some cases, the performance-enhancement effects of augmented feedback can be traced entirely to transformation of task-relevant information. Such a strategy is commonly employed in complex bimanual skills training, which often employ the Lissajous display. This technique involves turning a difficult to perceive, task-intrinsic perceptual variable (one which is necessary to learn to control) and consolidating it into an easier-to-perceive, unified Gestalt (Franz, Zelaznik, Swinnen, & Walter, 2001; Wilson, Snapp-Childs, Coats, et al., 2010). The outcome is that the learner no longer needs to monitor the activity of both hands, only the trace of a moving, on-screen dot. The task requirement then becomes to move the hands in such a way that the dot traces a recognisable shape on the display (depending on the required bimanual relationship, this may be a diagonal line, a circle, a crossed loop or other form, e.g. see: Kennedy, Wang, Panzer, & Shea, 2016). Complex bimanual coordination tasks can then be learned in minutes - rather than days, as is the case without feedback (Kovacs, Buchanan, & Shea, 2010). Note however that the task of controlling the dot bears little resemblance to the underlying motor task and its associated perceptual information[20]. The Lissajous display performs a substantial transformation of task-relevant information, and as a result, guidance effects are very pronounced (Maslovat et al., 2009; Ronsse et al., 2011). Learners have almost no ability to perform the task in the absence of this visual feedback after having trained with it. What may be required of new forms of feedback which aim to avoid the guidance effect is information which is not transformed from the basic, underlying kinematics of the motor task.

The Lissajous figure is an example of information-structural transformation which is very easy to recognise as such, because it clearly transforms information – from difficult-to-perceive information for the relationship between two separate hands into a single, easy-to-perceive, moving dot. However it is my contention that most, if not all visual feedback displays perform an information-structural transformation to some extent. Even a simple position-over-time display on a unidimensional reaching task is a transformation from the coordinate system of the hands to the coordinate system of the display (e.g. Vander Linden et al., 1993). The learner is required to learn the mapping - through which task goals and markers of their achievement are thereafter perceived. To a greater or lesser extent, the kinematics and intrinsic information associated with the actual underlying motor skill is lost. In the case of Lissajous feedback for bimanual coordination, the display completely subsumes the skill, such that controlling the display becomes the skill which is learned. It could be argued that the degree of guidance effect observed after withdrawal of feedback is a measure of the extent to which an information-structural transformation has occurred[21].

[20]Wilson, Snapp-Childs and Bingham(2010) show that the ability to detect and use the informational variable of 'bimanual relative phase' (a visual pattern which specifies the relationship between hands moving in space) underpins bimanual task performance without feedback. This information is not available in an interaction with a Lissajous display.

[21] Or to put it another way, the degree of a guidance effect observed is a measure of the extent to which the augmented feedback display has instantiated a different task than the one tested.

4.2.3 Sound and movement

There is some emerging evidence which suggests that sound/sonification as concurrent augmented feedback may not be subject to the guidance effect. Where a visual display might be seen as constituting a wholly different perceptual-motor task, it is possible that a sonic display might not. Instead, task-intrinsic information structure might be preserved in sonification, making feedback perceptually congruent with the underlying task kinematics, and allowing perception of both as a single perceptual-motor Gestalt. When sonified feedback is withdrawn, the dynamics of the task might not change substantially, thus perhaps allowing for successful maintenance of learned perception-action strategies. The current section reviews some of the theoretical and empirical basis for this suggestion.

The cross-disciplinary field of embodied cognition posits that sound and music perception are inseparable from the listener's physical and cultural context: that understanding through listening is intimately tied to particular ways of being in and interacting with the world (Antle et al., 2009; Roddy & Bridges, 2016; Schiavio & Altenmüller, 2015). Contemporary empirical research in Psychology shows the extent to which particular sounds are linked to action components of a learned task (Lahav, Saltzman, & Schlaug, 2007): sounds can influence coupled human movement, for example in gait (Wittwer, Webster, & Hill, 2013) and sensorimotor timing tasks (Rodger & Craig, 2014), and provide a listener with a great deal of information about the action-relevant properties of the sound-making event/agent (Gaver, 1993). In tasks which require movement with sound, agents tend to move in a way which reflects an implicit understanding of the real-world events that might produce the particular sonic pattern. With regard to music and melodies specifically, it has been argued that the events of musical motion are intuitively mapped onto the body of the listener, constraining patterns of bodily movement in line with the perceived musical ‘form’ (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2013; Repp, 1993; Rusconi et al., 2006).

This cross-modal relationship between sound and movement has been shown to hold with sonified movements as well - in cases where sonification is practiced in a 'direct' manner, that is, where sonification preserves dynamic features of task-intrinsic sensory stimulation (see section 3.1.5). Frid, Bresin, Alborno and Elblaus (2016) show that some structural commonalities between movement and sonification can be perceived by even mapping-naïve listeners, and that abstract drawings can reliably be paired with the sound patterns they represent. Similarly, Vinken et al. (2013) show that listeners can get an intuitive sense of the kind of motor task being performed (brushing teeth, pouring, stirring a pot, etc.) by listening to its sonified auditory profile. Schmitz et al. (2013) observed improved perceptual judgements about a swimmer's velocity when the swimmer's movements were sonified in a manner perceptually congruent with visual kinematics than when the accompanying sound information was incongruent, i.e. task-irrelevant. In some contexts, this link between movement and sound can be exploited through sonification to enhance performance in perception-action tasks. However it seems very likely that careful sonification mapping design is a critical factor in this effect, given the existence of some published examples of ineffective movement sonification. Sigrist, Rauter, Riener and Wolf (2013b) sonified multidimensional movement *error* in a rowing trajectory-matching task, and found no enhancement of learning. Using a similar, error-based mapping between current and target performance, Rosati, Oscari, Spagnol, Avanzini and Masiero (2012) found no benefit of sonification in a one-handed tracking task. It is possible that achieving cross-modal perceptual coherence between sonification and movement requires a direct mapping strategy, rather than sonification of error (for more on this idea, see section 3.1.7).

The implication of these findings (as they relate to the current study) is that sonification could be a form of augmented feedback which allows for concurrent pickup of task-intrinsic information - which is vital to avoid a guidance effect. In contrast with the structurally transformed information which is common in other forms of (mostly visual) feedback, sonification could share a variety of informational-structural and dynamic features with unmediated task-intrinsic information, allowing for joint perception of movement and feedback. This notion is similar to an approach advocated by Effenberg and his research group, who broadly study sonification in sporting contexts (Effenberg, 2005; Schmitz et al., 2013; Vinken et al., 2013). They propose that a strong structural ‘correlation’ between the perceived form of artificial sound and the kinetics/kinematics of the motor task is the key to harnessing multisensory integration processes and designing intuitive sonification systems. There is potential here for sound to become an auditory analogue of movement, perhaps even bridging the perceptual gap between augmented and intrinsic information (Johnson, 2007; Leman, 2008). However, controlled experiments to test the efficacy of sonification as augmented feedback for motor skill learning are still rare, and there is not yet any consensus on task performance in retention without feedback, i.e. the guidance effect (Sigrist et al., 2013a). Although most trials of sonification as feedback do not test motor performance in retention trials without feedback, some limited evidence does exist to support the ideas presented here. In rifle-shooting, long term benefits of sonified practice (up to 10 days post-practice) were found when the movements of the gun barrel were mapped to the pitch of a tone (Mononen, Viitasalo, Konttinen, & Era, 2003). In a more dramatic example, Ronsse et al. (2011) had two groups of participants learn a complex bimanual coordination task over the course of several days with visual (Lissajous) feedback and sonified feedback about hand position. Upon removal of feedback, a guidance effect was observed in the Lissajous group, but not in the sonification group.

4.2.4 The current study

If information provided by sonification is to share structural-dynamic features with intrinsic perceptual information, then movement-sound mapping design must be approached with great care. Sounds from the system must be perceivable alongside (and complement) intrinsic sensory information, and not be perceived as a separate to the actions which gave rise to them, otherwise a guidance effect may result. If sounds can be designed which appropriately represent movements of the learner, then more accurate performance – and maintenance of performance-benefits into no-feedback retention tests is to be expected. The current experiment aims to test some of these basic assumptions in a novel motor task.

This research has eventual application in movement rehabilitation following stroke (Rosati et al., 2013; Scholz, Rhode, Großbach, Rollnik, & Altenmüller, 2015). Much of the published research on stroke rehabilitation involves training functional skills incorporating many biomechanical degrees of freedom, such as reaching for small objects, manipulating them in some way and placing them at a new location (Dobkin, 2004; Nudo, 2006). This is meant to approximate real-life tasks such as making dinner and drinks, with the aim that training in a slightly constrained, artificial (laboratory) scenario will transfer to daily life. Several measures of general performance fluency are normally taken, including global measures like time taken to complete the task, and slightly more domain-specific measures such as movement smoothness, which is often lacking following stroke. In light of this literature, a novel motor task was designed which incorporated many of the same features as the kind of training involved in stroke rehabilitation, scaled up in difficulty for a healthy sample.

4.3 Method

Participants

45 participants were recruited from the local Psychology undergraduate population and staff in the department. Undergraduates received partial course credit for their participation. Left-handed participants were excluded from participation at the recruitment stage. All participants reported normal hearing and upper-body mobility. No participants recruited were professional musicians.

Informed consent was obtained from all individual participants included in the study. Ethical approval for this study was granted by the School of Psychology Ethics Board at Queen’s University, Belfast.

Motor task

In this task, participants learned to perform a timed sequence of manipulations of small objects while seated at a table. This is intended to mimic the kind of complex procedural motor learning common in everyday life that can be difficult to master. For example, driving a car requires very carefully timed and ordered manipulations of a variety of control devices (wheel, gearstick, clutch, brake and accelerator) whist simultaneously maintaining a high level of situational and geographical awareness. This task can be quite difficult for beginners, but becomes easier with repeated practice (Shinar, Meir, & Ben-Shoham, 1998). Keeping this in mind, and taking some inspiration from the stroke rehabilitation literature (Dobkin, 2004), an object manipulation task with a focus on careful ordering and timing was designed. Participants were required to undertake extensive practice to master the task. They learned to reproduce a short series of movements, in order, and at a predetermined correct speed. Temporal and spatial accuracy was recorded over the course of practice using optical motion capture.

The three objects used were selected from a set of plastic, three-dimensional shapes (Learning Resources Ltd.). Shapes were of equal height when positioned upright and comprised a yellow triangular prism, a red square-ended cuboid and a blue cylinder (see Figures 4.1 and 4.2). Shapes were chosen from the set by the experimenter to be as subjectively similar in 'grasp-ability' as possible. A set of reflective spherical markers were attached to the top surface of each shape. This allowed them to be independently identifiable by the software used to run the experiment.

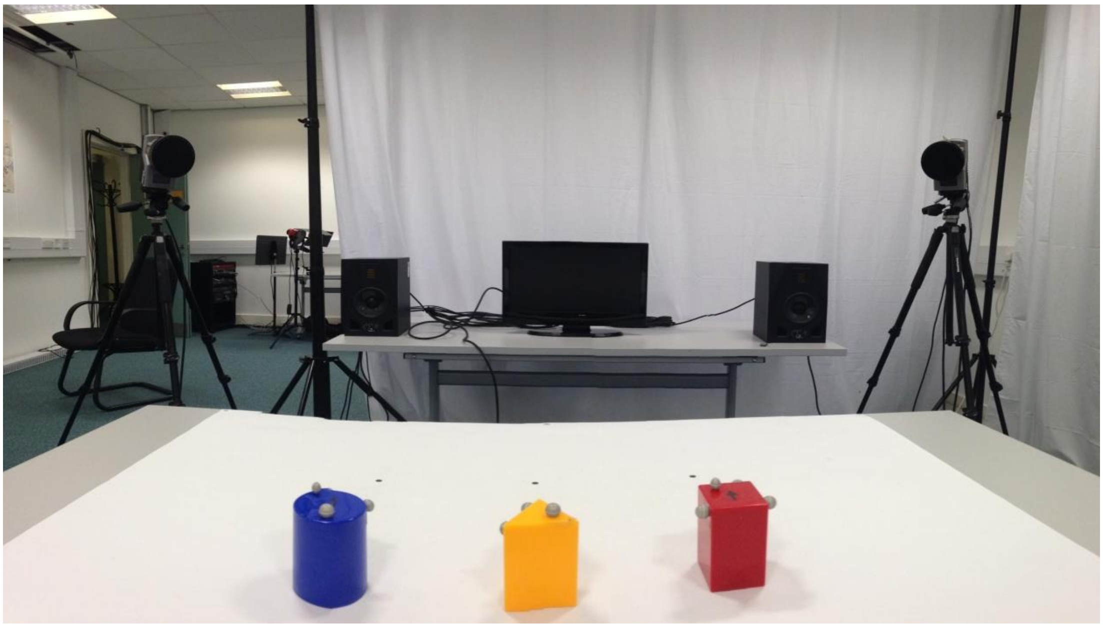

Figure 4.1: Experimental apparatus: shapes, screen, speakers and two of four motion-capture cameras as seen from the participant's perspective.

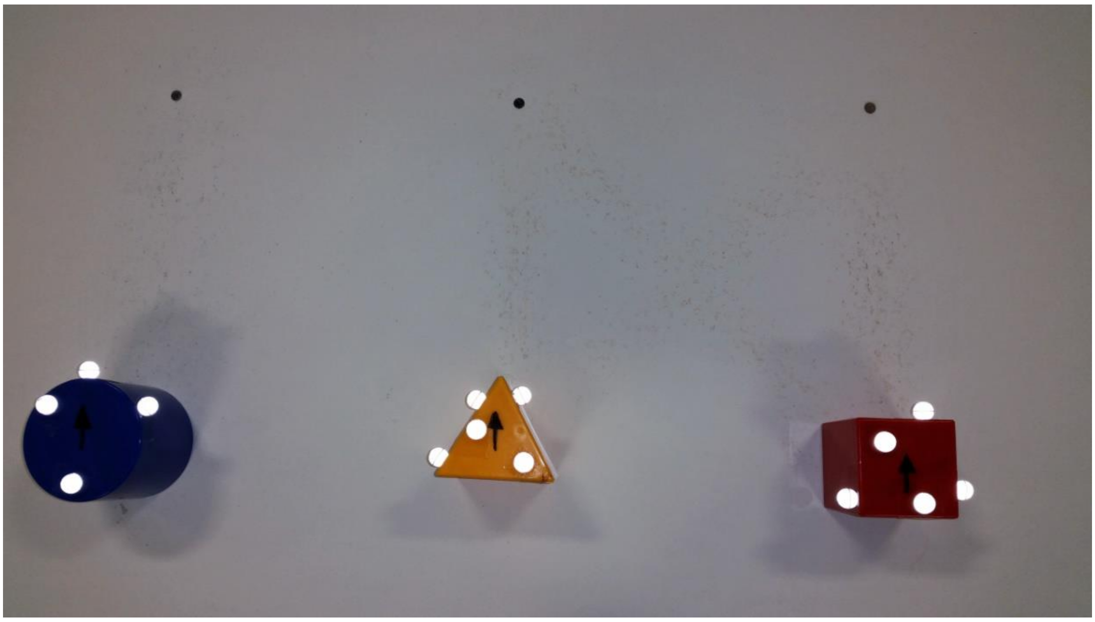

Figure 4.2: Experimental apparatus as seen from above, with reflective markers highlighted.

The task to be learned required participants to move the three shapes between six target zones on a tabletop workspace, each defined by a circular black sticker of 1cm diameter. The task sequence can be seen in the accompanying schematic (Figure 4.3). A link to a video of task performance (with sonification) can be found in Appendix A. Participants were presented with an animated demo, programmed in *Processing*[22]. The goal, as explained to participants, was to move the shapes in the ordered sequence shown in the demo, within exactly 8.3 seconds (the length of the demo), whilst keeping the movement of the cylinder shape as smooth (i.e. non-jerky) as possible, following the same movement profile shown in the demo (which had a sinusoidal velocity profile). An additional requirement was to ensure that the cylinder always arrived at a target zone at the same time as the other shape in motion at that time. The demo animation was displayed on a 17-inch monitor, positioned approximately 1.5 m in front of the participant (see Figure 4.1).

[22] The 2-dimensional demo showed the three shapes, seen from directly above, moving around the workspace between target zones as shown in the task schematic.

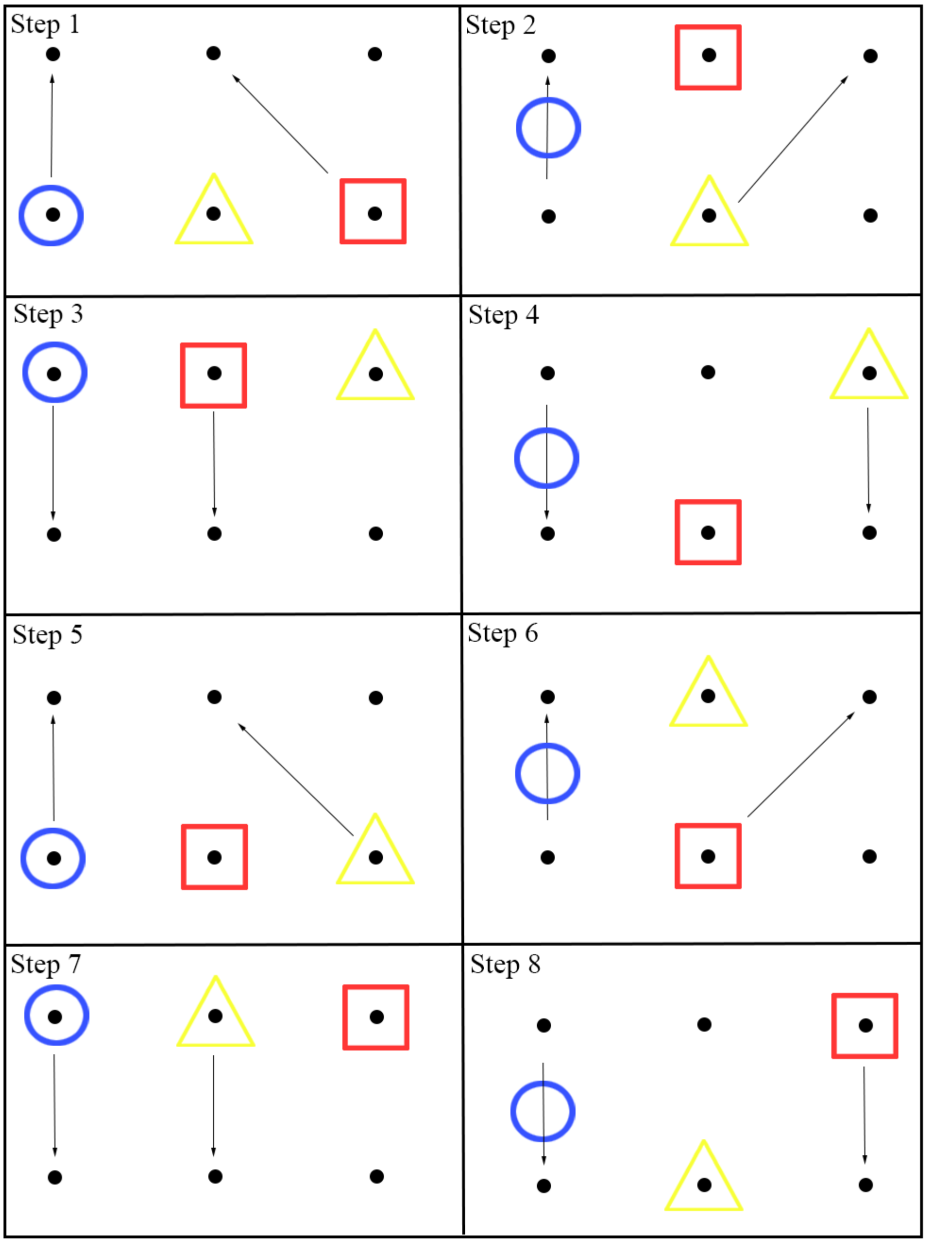

Figure 4.3: Task schematic. The left hand held and controlled the cylinder at all times. Two movements of the prism and cuboid shapes occurred for each movement of the cylinder. Shapes began and finished at the same positions. The animated demo was visually similar to this schematic

Optical motion-capture

Movements of the shapes were tracked over time using four optical motion capture cameras (Qualisys) capturing at 300Hz. Cartesian coordinate data were streamed at 20 Hz to a patch written in *Max/MSP 5.9* for use in live sonification. Captured data were additionally saved at full temporal resolution for later movement analysis in *MatLab*.

Sonification of movement

Sound was played to participants during the experiment through a pair of speakers positioned 1.5 metres in front of the workspace, approximately 30-degrees anticlockwise (for the left speaker) and 30-degrees clockwise (for the right speaker) from the participant's forward-facing position. See figure 4.1 for the arrangement of the lab.

Sonification of movement was designed by the experimenter using an iterative procedure. A few main principles guided the design process:

The sounds produced by movement of the shapes should perceptually 'fit' with the movements themselves. This was achieved by adhering roughly to sonification design principles devised by the Auditory Display community which are inspired by Ecological Psychoacoustics (Gaver, 1993; Saue, 2000; B. N. Walker & Kramer, 2004) and the idea of a 'natural control mapping' in product design (D. Norman, 2002). In practice, this meant that there was only audible sound when movement occurred, sound quality reflected the amount of energy imparted by the mover, and certain ecological mapping metaphors were applied (described below). The sounds produced by each shape were given perceptually discrete 'identities' so that activity of each shape could be discerned. This is again inspired by discussions of perceptual streaming in Auditory Psychoacoustics and Auditory Display (Barrass & Vickers, 2011; Bregman, 1990; Flowers, 2005).

The iterative design process was driven by subjective feedback and impressions gleaned through testing by the experimenter and colleagues in both the School of Psychology and the Sonic Arts Research Centre at QUB. The aim was to arrive at movement-sonification mappings which intuitively made sense to a moving individual in the context of the task motions.

All three shapes triggered an enveloped burst of filtered pink noise when moved > 1.5cm in any direction. Movement trajectories thus produced a pattern of granular sound, the density of which was directly mapped to velocity. This mapping was inspired by a metaphor of an object moving over corrugated metal. It is possible to infer the speed of said object from sound alone if the space between ridges is known. The shape of the envelope and filter was varied between the shapes in order to create distinct timbres of sound for each. The cylinder had a characteristic ‘rich’, but electronic sound. The prism sounded ‘cleaner’ and ‘lighter’ in comparison. The cuboid sounded like a percussive ‘clack’, with a subtle background hiss; a colleague described it as akin to a medieval winch. The central frequency of the pink noise burst was varied with distance from the participant. Pitch increased as the shapes were brought closer to the participant and decreased as they were pushed away. This mapping was designed in keeping with well-known principles of Ecological Psychoacoustics. To elaborate, more spectral content in the high-frequency end is audible from a sounding object the closer it is to the listener, and moving-sounding objects sound as if their pitch is increasing as they come closer to the listener – known as the Doppler effect (see: Gaver, 1993). Related research in Auditory Display indicates that the use of familiar, real-world action-sound mappings (or analogical mappings) can help scaffold listener understanding of otherwise abstract sonic information (Rath & Schleicher, 2008; Roddy & Bridges, 2016; Walker & Kramer, 2005).

A shape's arrival at any of the six target zones was denoted by the synthesised sound of a bass drum with a fast decay. To achieve this, the position of a single, central marker on each shape was tracked, and triggered a beat event when it entered a zone (defined in the x and y axes, 4 by 4cm) centred on each target. It was intended that the regular structure of these beats in both the demo and the workspace would enable participants to more easily keep in time. The drum was triggered when a shape came within a predefined range around a target zone. Each shape's location on the horizontal plane was mapped to stereo balance between two speakers, such that sound location corresponded to shape location. Loudness of beats was modulated by distance from the participant - shape arrivals at the top of the workspace were slightly quieter than arrivals at the closer side.

Experimental conditions

Two sonification conditions and one silent control condition were included in this experiment. Participants were pseudorandomly assigned to one of the three independent conditions. In the first sonification condition (hereafter referred to as “Full Sonification”), all movement events were sonified. This included the full trajectory of shape movement across the workspace (via enveloped bursts of pink noise as described above) and the moment of arrival at a target zone (via a triggered bass drum sounds). In the second sonification condition (hereafter referred to as “Arrival Sonification”), only arrivals of the shapes at target zones was sonified. This sonification was provided in an identical manner to the arrival sonification used in the Full Sonification condition. Arrival Sonification was included to test the efficacy of a more reduced style of sonification purely focussed on the provision and prescription (via the demo) of discrete-event timing information. No sound was provided in the control condition.

Practice and testing procedure

Before the practice phase, all participants were given five minutes to familiarise themselves with the workspace and the basic requirements of the task. Participants were encouraged to move the shapes around the workspace in straight lines between target zones. During this time, sonification of movement was switched on for participants in the Full Sonification and Arrival Sonification conditions. A familiarisation period has been shown to be necessary for participant understanding of multidimensional movement sonification systems (Schaffert & Mattes, 2015; Sigrist et al., 2011).

During this pre-practice phase, all participants were shown the demo animation two times and asked to reproduce the ordered sequence of movements. The demo animation was silent in this phase. If a given participant was able to reproduce the order of the sequence of movements correctly, they proceeded to the practice phase. If not, they were allowed to view the demo an additional time.

Practice consisted of 100 trials, split across two 1-hour sessions, a maximum of four days apart. Each trial commenced with a single play of the demo. The demo was sonified according to the condition assignment of the participant. I.e. those in the control condition watched a silent demo, those in the Full Sonification condition watched with trajectory and target arrival sonification and those in the Arrival Sonification condition watched with only arrival sonification. Immediately following the demo, participants attempted to perform the same set of movements, matching the spatial and temporal components of the animation. Perfect performance of the task would result in exactly the same sound produced by the workspace as heard in the demo. Participants were informed that they should aim to replicate both the visuospatial and auditory (where appropriate) elements of the demo animation.

The time between the commencement of movement and the final cessation of movement was calculated and presented to participants verbally as terminal feedback on each trial. This was intended to provide some knowledge-of-results feedback which could be motivating and encourage engagement with the task. The target time was 8.3 seconds.

Trials were conducted in blocks of ten, with a short break between blocks. During this break, participants were reminded of the goals of the task:

1) To match the demo (visuals and sound)
2) To complete the sequence in 8.3 seconds
3) To move the cylinder as smoothly as possible (with minimal jerkiness)
4) To time the arrival of the cylinder to co-occur with the arrival of another shape at its target zone (see Figure 4.3)

At the end of the second session, when 100 practice trials had been completed, a five-minute break was scheduled. During this time, participants did not interact with the workspace in any way. All participants then completed three retention trials with no demo, sonification or terminal feedback.

After three days, participants returned for another retention session (three trials, performed under the same conditions as described previously). A final retention session occurred exactly one week after the second practice session.

Analysis

A range of primary performance measures were taken to produce, in combination, a general picture of task fluency. These measures reflected the task instructions given to participants.

The main measure of performance was global timing error. Participants were required to complete the task in exactly the same time as displayed in the demo animation (8.3 s). Trial times were measured as the time between the beginning of the first shape movement and the end of the last. The absolute difference between achieved trial time and ideal trial time was calculated for each trial to give the global timing error measure.

The velocity of the cylinder shape was perfectly sinusoidal in the demo animation; participants were expected to try to match this movement style. Harmonicity of participant movement of the cylinder was quantified by correlating acceleration and position within each trial to extract an r value for each trial. With perfectly sinusoidal cylinder movement, there should be an r value of 1. These values then underwent a Fisher z-transformation to correct for possible skewing and to allow for parametric statistical analyses.

The timing of the arrival of the cylinder shape at its target destination and that of whichever other shape is moving at the same time is identical in the demo. To measure the extent to which this was achieved by participants, the time of every right-hand arrival was recorded. Those arrival events which co-occur with cylinder arrivals in the demo (that is, every second RH shape arrival) were compared to the corresponding cylinder arrival to obtain a difference score (in seconds). This procedure produced four values per trial, which were averaged to produce a per-trial measure of intermanual timing. Lower values reflect better synchronisation performance.

Per-trial measures for practice and retention were averaged together for analysis as discrete blocks of trials, per participant (i.e. trials 1-10, 11-20, etc.). For each measure, this produced ten scores for each participant during the practice phase and three each in retention.

Mixed ANOVA were employed to test for main effects and interactions. To detect effects of sonification condition assignment on performance during practice, planned comparisons were performed for each measure at the first and last practice trial. Where the assumption of sphericity is violated, Greenhouse-Geisser corrected values are reported. Where appropriate, direct comparisons between group performance were performed using *t*-tests. Learning curves produced during the practice stage are investigated with the use of regression.

The data for one participant in the control condition at practice block 2 were unavailable due to camera error. These data were omitted from the analysis and not replaced.

4.4 Results

4.4.1 Global timing error

Mean timing error scores for each group across the different phases of the experiment are shown in Figure 4.4. A mixed ANOVA on practice blocks 1-10 with feedback condition as a between-subjects factor and practice block as a within-subjects factor revealed a significant main effect of feedback condition: $F(2, 41) = 5.388$, $p = .008$, $\eta^2 = .053$, practice block: $F(1.662, 68.135) = 87.026$, $p < .001$, $\eta^2 = .461$, and an interaction between feedback condition and practice block: $F(3.324, 68.135) = 6.218$, $p = .001$, $\eta^2 = .065$. Planned comparisons were performed between feedback conditions at blocks 1 and 10 ($\alpha = .008$; Bonferroni correction for six comparisons), revealing significant score differences at the first practice block between Full Sonification (M = 6.63 s) and Control (M = 2.94 s): $p = .001$, Cohen's $d = 1.452$), but not between Full Sonification and Arrival Sonification (M = 5.55 s): $p = .279$. The score difference between Arrival Sonification and Control at block one did not reach the corrected criterion for significance ($p = .012$). At block 10, no significant differences were observed between Full Sonification and Control: $p = .487$; Full Sonification and Arrival Sonification: $p = .619$; nor between Arrival Sonification and Control: $p = .842$. This indicates that performance had converged over practice blocks.

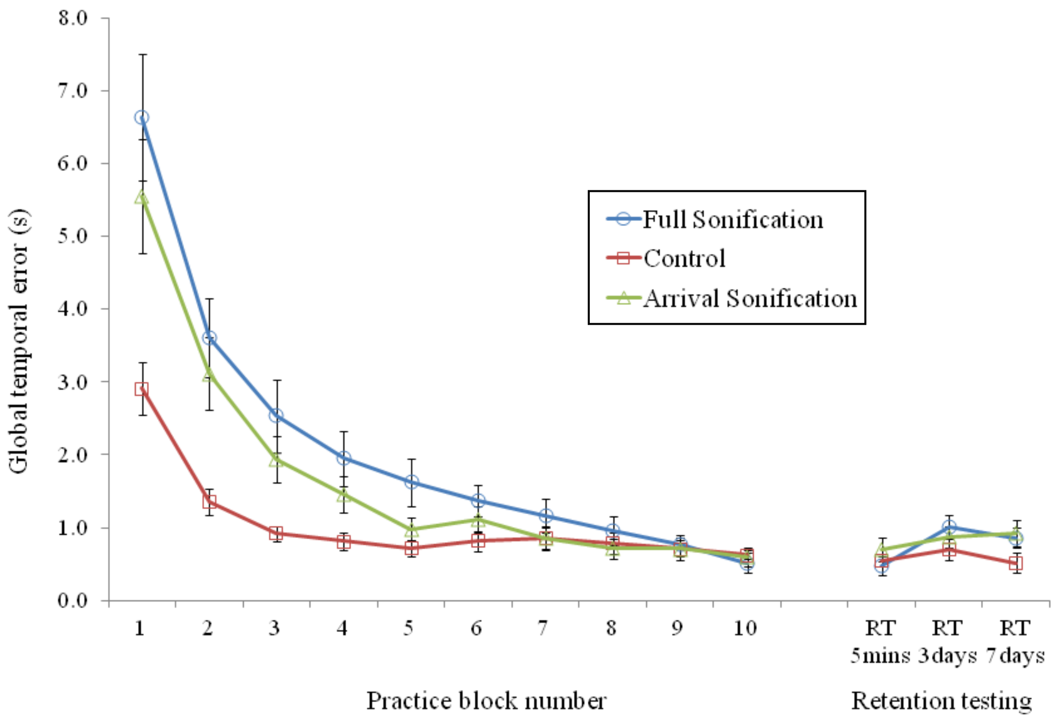

Figure 4.4: Absolute global temporal error over the course of practice and retention testing for three experimental conditions. Error bars are standard error.

To compare group performance in retention, a mixed ANOVA was run on data obtained from all three retention-testing occasions with feedback condition as a between-subjects factor and retention block as a within-subjects factor. A significant main effect of block was detected $F(2, 84) = 6.475$, $p = .002$, $\eta^2 = .049$, but no main effect of condition: $F(2, 42) = 1.091$, $p = .345$ and no interaction between block and condition: $F(2, 84) = 1.831$, $p = .130$. Pairwise comparisons between testing blocks (collapsed across conditions; $\alpha = .017$) showed that timing error was significantly greater after 3 days as compared to 5 minutes post-practice: $t(44) = -3.773$, $p < .001$, Cohen's $d = .729$. Differences in error between 5 minutes and 7 days post-practice and between 3 days and 7 days post-practice were not significant ($p = .028$; $.255$ respectively).

4.4.2 Smoothness/harmonicity of cylinder movement

Mean smoothness scores for the three groups over the different phases of the experiment are shown in Figure 4.5. A mixed ANOVA with feedback condition as a between-subjects factor and practice block as a within-subjects factor was performed on acquisition data (blocks 1-10) to investigate differential effects of sonified feedback during learning. ANOVA detected no significant main effect of condition: $F(2, 41) = 2.654$, $p = .082$; a significant main effect of block: $F(2.822, 115.690) = 35.708$, $p < .001$, $\eta^2 = .184$; and a significant interaction between block and condition: $F(5.643, 115.690) = 1.891$, $p = .010$, $\eta^2 = .020$. Planned comparisons were performed between feedback conditions at blocks 1 and 10 ($\alpha = .008$; Bonferroni correction for six comparisons). This revealed a significant difference in scores at block 1 between the Control condition (M = .392, SD = .070) and Full Sonification (M = .298, SD = .058): $p < .001$, Cohen's $d = 1.462$, but not between the Control condition and Arrival Sonification (M = .322, SD = .089): $p = .012$. No significant between-groups differences were observed at block 10. Results were as follows: between Full Sonification and Control, $p = .935$; between Full Sonification and Arrival Sonification, $p = .135$; and between Arrival Sonification and Control, $p = .205$. These analyses indicate that an early advantage in the control condition was attenuated throughout practice, and that on the measure of smoothness, intergroup performance converged over time.

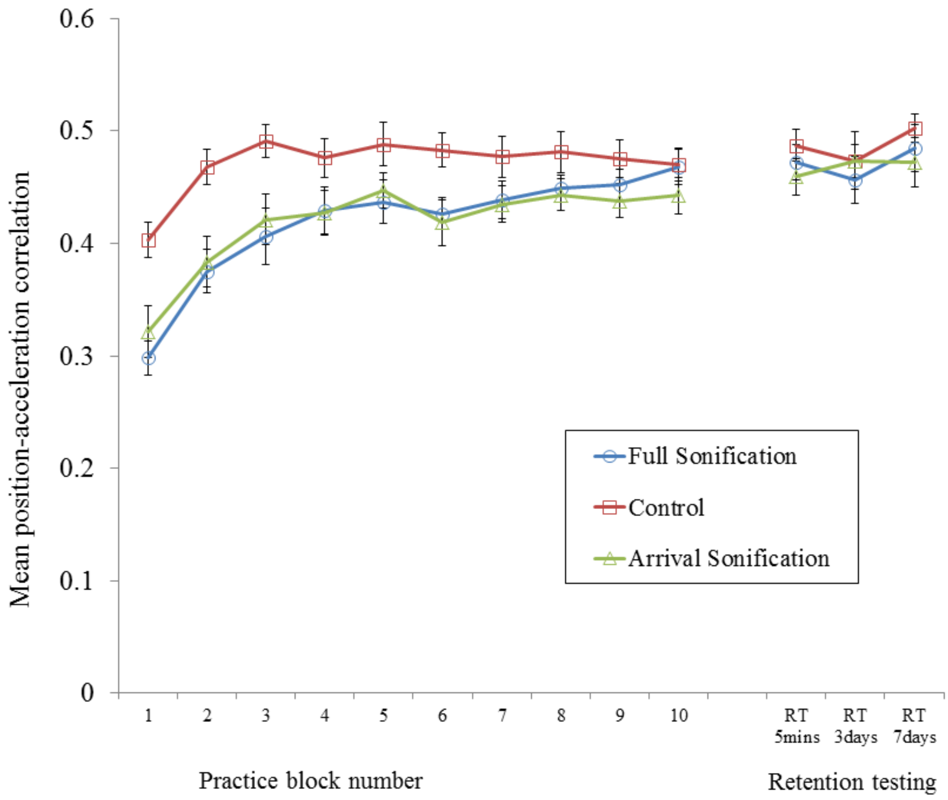

Figure 4.5: Correlation *r* values, Fisher z-transformed (representing smoothness of cylinder movement) for practice and retention. Higher values denote movements which more closely match the demo. Error bars are standard error.

Intergroup performance in retention testing over the next several days was examined by submitting harmonicity data to a mixed ANOVA with feedback condition as a between-subjects factor and testing block (5 minutes, 3 days and 7 days) as a within-subjects factor. This detected no main effect of feedback condition: $F(2, 42) = .207$, $p = .814$, a significant main effect of testing block: $F(1.691, 71.004) = 3.642$, $p = .038$, $\eta^2 = .021$, and no significant interaction between feedback condition and testing block: $F(3.381, 71.004) = .681$, $p = .584$. Pairwise comparisons were performed between testing blocks, collapsed across condition ($\alpha = .016$). These revealed a significant difference between smoothness scores at the first retention test (5 mins post-practice) and the last (7 days post-practice), with performance superior at the last: $p = .012$. Scores did not differ significantly between the first and second (3 days post-practice) retention tests: p = .496, nor between the second and last: p = .027.

4.4.3 Intermanual timing error

Mean intermanual timing error for the three conditions across all phases of the experiment can be seen in Figure 6. A mixed ANOVA with condition as a between-subjects factor and practice block as a within-subjects factor was performed on acquisition data from blocks 1-10. A significant main effect of feedback condition was not found: F(2, 41) = .908, p = .411. A significant main effect of practice block was found: F(1.458, 59.887) = 22.425, p < .001, η^2 = .266, but no significant interaction between practice block and feedback condition: F(2.916, 59.778) = 1.130, p = .343. As feedback condition appeared not to have an effect on intermanual timing error in acquisition, no further analyses were performed on acquisition data for this measure.

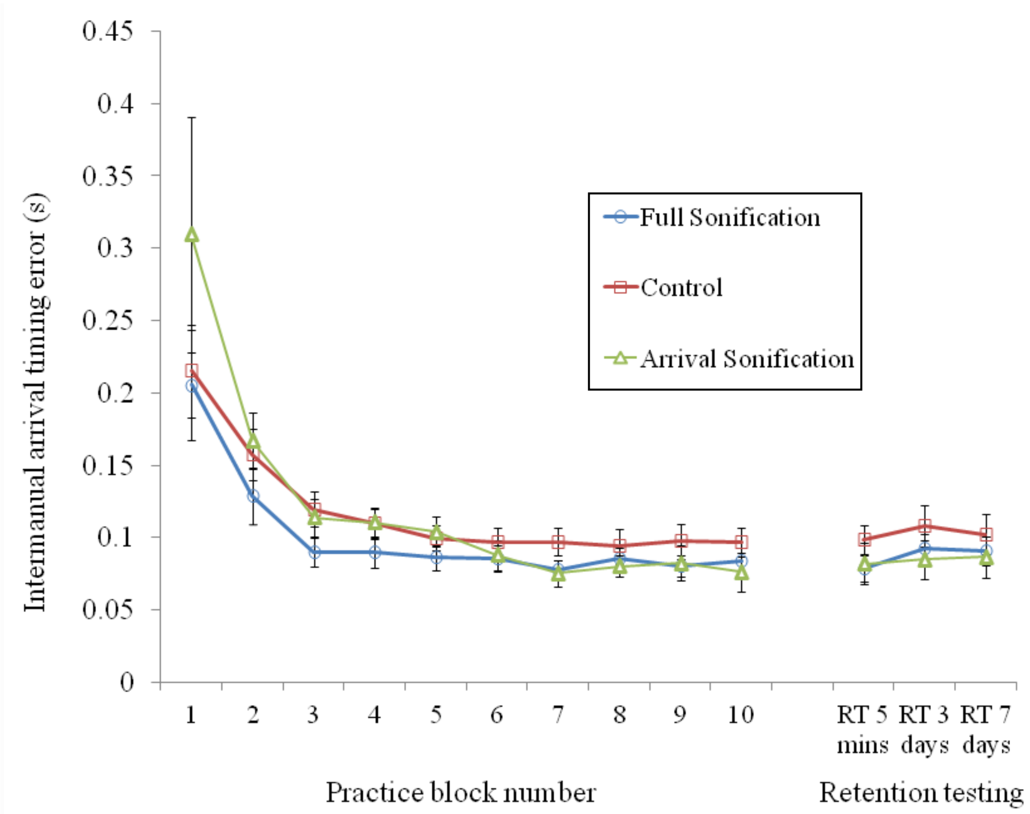

Figure 4.6: Mean time difference (s) between cylinder arrival and the arrival of the contemporaneous right-hand shape. Lower values denote better synchronicity. Error bars are standard error.

A mixed ANOVA was employed to test for an effect of feedback condition or time delay in retention. ANOVA revealed no significant main effect of feedback condition: $F(2, 42) = .804$, $p = .454$, or testing block: $F(2, 84) = 1.482$, $p = .233$, and no significant interaction between feedback condition and testing block: $F(4, 84) = .232$, $p = .919$.

4.4.4 Learning curve analysis

Exponential curves were fitted to individual participants' acquisition data to examine rate of learning for each performance variable using an iterative procedure in MatLab. Best-fitting curves were ensured by performing the fitting procedure 200 times per participant (for a given variable), and selecting the curve with minimum fitting error (from the 200 generated). This yielded an asymptote, range and slope parameter for each participant, thus enabling group-level comparisons of learning curves. Holm-Bonferroni-corrected *p* values are reported for significance tests in this section. Results are reported for each variable, in order.

For global temporal error, in the Control condition, the mean slope value was 1.1666 (S.D. = .806). A one-tailed *t* test (against 0) revealed that curve slopes were significantly greater than zero: $t(14) = 5.607$; $p < .015$. In the Full Sonification condition, the mean slope value was .803 (S.D. = .50). A one-tailed *t* test (against 0) revealed that curve slopes were significantly greater than zero: $t(14) = 6.244$, $p < .015$. In the Arrival Sonification condition, the mean slope value was .735 (S.D. = .423). A one-tailed *t* test (against 0) revealed that curve slopes were significantly greater than zero: $t(14) = 6.756$, $p < .015$. To compare learning curves between experimental conditions, slope data were submitted to a one-way ANOVA, which revealed no significant effect of condition on slope: $F(2, 44) = 2.252$, $p = .354$.

For cylinder smoothness, in the Control condition, the mean slope value was 5.267 (S.D. = 4.946). A one-tailed *t* test (against 0) revealed that curve slopes were significantly greater than zero: $t(14) = 4.124$, $p < .015$. In the Full Sonification condition, the mean slope value was 1.269 (S.D. = 3.037) and slopes were not significantly greater than zero: $t(14) = 1.618$, $p = .256$. In the Arrival Sonification condition, the mean slope value was 1.278 (S.D. = 2.605) and slopes were not significantly greater than zero: $t(14) = 1.900$, $p = .195$. A one-way ANOVA was performed to compare slope values between experimental conditions: $F(2, 44) = 5.911$, $p = .055$, revealing no effect of condition on slope.

For intermanual arrival timing error, in the control condition, the mean slope value was 2.342 (S.D. = 3.500). A one-tailed *t* test (against 0) revealed that curve slopes were not significantly greater than zero: $t(14) = 2.467$, $p = .084$. In the Full Sonification condition, the mean slope value was 3.586 (S.D. = 5.580). A one-tailed *t* test (against 0) revealed that curve slopes were not significantly greater than zero: $t(14) = 3.033$, $p = .055$. In the Arrival Sonification condition, the mean slope value was 1.280 (S.D. = 1.958). A one-tailed *t* test (against 0) revealed that curve slopes were not significantly greater than zero: $t(14) = 2.532$, $p = .084$. To compare learning curves between experimental conditions, slope data were submitted to a one-way ANOVA, which revealed no significant effect of condition on slope: $F(2, 44) = 1.534$, $p = .456$.

4.4.5 Further analysis

An additional exploratory analysis was performed to investigate a possible explanation for the frequently observed poorer performance in both sonification conditions relative to the control condition. It was hypothesised that the use of sonification may impose constraints on movement performance which do not exist in the control condition. More specifically, the requirement to produce sound congruent with the demonstration may have encouraged the development of movement strategies to control sound at the expense of the measures of performance reported so far. In all conditions, participants must place each shape on an arrival marker at the end of a movement trajectory. However only in the sonification conditions is this marker audibly defined as a 'zone' - an arrival within which produces feedback - a beat. Although arrival zones were generously sized, the existence of a zone, paired with feedback related to accuracy, may have encouraged a strategy of positional accuracy-monitoring in the sonification conditions. In the control condition, participants were not constrained by the need to hear affirmative feedback about accuracy, and so may have attended more to achieving the stated goals of the task: overall timing, smoothness of cylinder movement and intermanual timing. To investigate this possibility, the Euclidian distance was taken between the position of the first-moved right-hand shape (after it had come to a stop) and the exact centre of the appropriate target zone. The resulting error scores (in millimetres) for each condition, over time, are shown in Figure 4.7. The first movement arrival was selected for analysis on the basis that it represents a particularly salient perceptual-motor event - the first sub-task completed and first piece of audible feedback in each new trial. Therefore it might be the most likely place to see fine positional control in the sonification conditions.

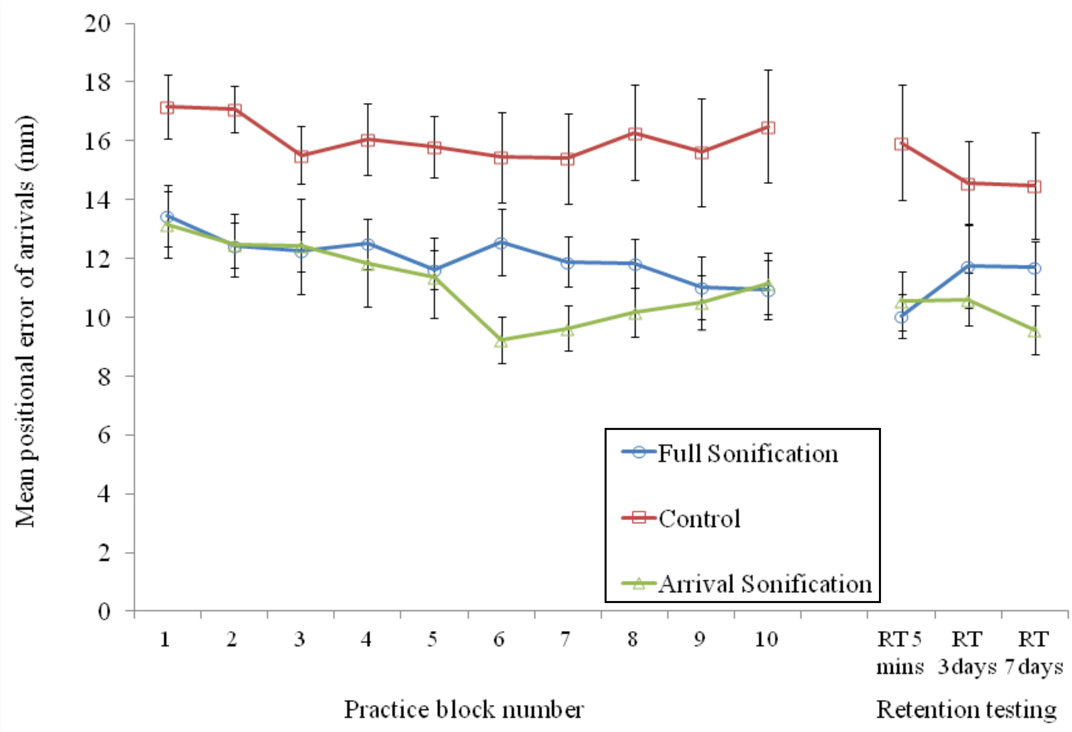

Figure 4.7: Mean positional error of arrivals/stops on the first movement of each trial for three feedback conditions across practice and retention testing. Lower values denote greater accuracy. Error bars are standard error.

To test for effects of feedback and practice during the acquisition phase, positional-error data from practice blocks 1-10 were submitted to a mixed ANOVA with practice block as a within-subjects factor and feedback condition as a between-subjects factor. This revealed a significant main effect of feedback condition: $F(2, 42) = 9.034$, $p = 0.001$, $\eta^2 = .189$. There was no significant effect of block number: $F(3.080, 129.361) = 2.178$, $p = .092$.

Planned comparisons were performed between feedback conditions at blocks 1 and 10 ($\alpha = .008$; Bonferroni correction for six comparisons). At practice block 1, shape arrivals in the Arrival Sonification condition (M = 13.15, SD = 4.34) were not significantly more accurate than in the Control condition: (M = 17.08, SD = 3.90): $p = .012$. Neither was the mean difference between the Full Sonification condition (mean positional error = 13.40 mm, SD = 4.03) and the Control condition: $p = .018$. There was no significant difference detected between the two sonification conditions: $p = .866$. At block 10, participants displayed significantly lower error scores in the Full Sonification condition (M = 11.20, SD = 3.96) and the Arrival Sonification condition (M = 11.15, SD = 4.12) than the Control condition (M = 16.44, SD = 6.93): $p = .008$; .008, Cohen's $d = .929$; .927 for each comparison respectively. A significant difference between the two sonification conditions was not detected: $p = .866$.

To test performance in retention, a mixed ANOVA was employed. A significant main effect of feedback condition was detected: $F(2, 42) = 6.828$, $p = .003$, $\eta^2 = .245$. No effect of testing block was detected: $F(2, 84) = .118$, $p = .889$, nor an interaction between condition and testing block: $F(4, 84) = .982$, $p = .422$. Pairwise comparisons were performed between feedback conditions, collapsed across testing blocks ($\alpha = .017$; Bonferroni correction for 3 comparisons). Positional error was revealed to be significantly lower in the Full Sonification condition than Control (p = .007), and lower in the Arrival Sonification than Control (p = .001). No Significant difference between the two sonification conditions was observed (p = .535).

From visual inspection of learning curves in Figure 4.4, it appears as though an asymptote was not reached in the Full Sonification condition, unlike in the Control and Arrival Sonification conditions. This raises the interesting possibility that learning was still ongoing at the close of practice for participants in the Full Sonification condition. Based on the shape of learning curves for individual participants, it is possible to estimate an asymptote value for the Full Sonification condition, using the iterative curve-fitting procedure described in section 4.4.4. This approach theoretically extrapolates the existing learning curve to provide a global temporal error score at which we might expect to see performance reach asymptote if more practice trials had been performed. Exponential curve-fitting revealed that although the mean asymptote value for temporal error was indeed lower in the Full Sonification condition (.561s) than the Arrival Sonification and Control conditions (.636s and .817s respectively), there was no significant effect of condition on asymptote: $F(2, 44) = .642$, $p = .531$.

4.5 Discussion

The present study was intended to address an open question in the feedback literature, whether sonification as concurrent augmented feedback is subject to the frequently-reported 'guidance effect' (Anderson et al., 2005). When visual feedback is delivered concurrently with movement, improved performance is usually observed, at the price of dependence on feedback - performance declines when feedback is removed (Maslovat et al., 2009; Ronsse et al., 2011; Schmidt et al., 1989). It is as yet unclear whether sonification can overcome this effect as a general rule. It was suggested here that a mapping strategy which preserves the spatiotemporal and dynamic features of the underlying motor task might allow for pickup of task-intrinsic information alongside augmented feedback. Significant slope values for modelled learning curves observed for several performance variables indicate that all three experimental groups showed improvement on a variety of task measures over the course of practice. However a benefit of sonification did not emerge for the main performance measures in this task.

4.5.1 Sonification as a hindrance

For global temporal error, there appeared to be an initial advantage for participants practicing without sound relative to Full Sonification. A mean difference in the same direction was found between Control and Arrival Sonification, but did not reach the corrected criterion for significance. By the end of the practice phase, performance in all conditions improved and between-groups differences were attenuated completely. A slight performance decline in medium-term retention was observed in the Full Sonification condition relative to control. A similar lack of advantage of sonification of either kind was observed in for the measure of intermanual timing error. This observed lack of performance enhancement (and arguably, evidence of hindrance during acquisition) with the provision of movement sonification represents an unexpected finding. The sensorimotor timing literature shows that actions coupled with sound can be performed with less temporal error than silent actions (Repp, 2005; Rodger & Craig, 2014), and it was hypothesised that movement sonification could confer a similar advantage here. There are several possibilities which could explain this pattern of results.

The current task, speeded bimanual object manipulation, was designed to mimic the kinds of real-world motor skills which are highly valued in performance assessment and training after certain kinds of stroke, i.e. those which result in motor impairment (Dobkin, 2004). Reaching, manipulating and moving handheld objects are fundamental, everyday skills, performance of which underpins many common motor tasks (cooking, driving etc.). The current task was scaled up in difficulty with the aim of mimicking the basic motor skills, while presenting a challenge to physically unimpaired participants. However, speeded performance of the task, together with the additional task requirements (smooth movement, bimanual arrival synchronicity) may in fact have presented *too much* of a challenge, overshadowing any timing advantage which might otherwise have been conferred through movement sonification. The boost in sensorimotor timing performance commonly reported in the literature is most-frequently reported in simpler tasks, e.g. tapping on a force plate (Repp & Penel, 2002). In a broad review of motor skill learning literature, Wulf and Shea (2002) caution that results obtained from the study of simple actions may not always scale up to more complex motor skills, especially where augmented feedback is employed. They argue that the reduced motor-planning demand typical of simpler tasks (e.g. tapping) leaves more space in attention for the pickup and use of augmented information, whereas the multiple competing demands of complex tasks (such as this) do not – and complex task performance could actually suffer as a result of information overload. The Arrival Sonification condition was included partly in acknowledgement of this possibility; participants should have less difficulty picking up task-relevant timing information when the necessary sonic information cannot be obscured by the sound of another shape's trajectory. However again, timing performance (on either temporal measure) was not improved by the availability of Arrival Sonification. It is worth noting that the 'information overload' described by Wulf and Shea (2002) concerns not only the amount of sensory feedback given/available, but the overall complexity of the task; the number of biomechanical degrees of freedom involved and the number of individual task requirements are suggested to have a non-linear, cumulative effect on movement planning demand. Given the highly complex nature of the current task, it is plausible that extra demands imposed by sonification (i.e. the requirement to match the sonic profile of the demonstration as well as the movements) created a task which was more difficult - rather than easier, as was intended.

Broadly similar results were observed for the measure of cylinder harmonicity. Harmonic or 'pendular' motion is characteristic of healthy reaches, but is difficult to produce under conditions of speeded movement or when aiming for a small target (Bootsma, Fernandez, & Mottet, 2004). It was expected that performance on this measure would be most effective in the Full Sonification condition, as shape velocity over time was perceivable via the granular density of filtered pink noise bursts produced by movement of the cylinder. In the demo animation, participants had access not only to a visual, but auditory model of how cylinder movement should be performed; the task was to match both components. However, sonification of either type led to less effective performance in early acquisition relative to control. Performance between groups was equalised through practice, and no advantage of sonification emerged in practice or retention.

4.5.2 The guidance effect

The current results leave interpretation in light of the guidance effect unclear. In retention, no decline in performance relative to control was observed after practicing with sonification. The guidance hypothesis predicts that feedback delivered on all practice trials will cause dependence, and a performance decline upon its withdrawal (Maslovat et al., 2009; Schmidt et al., 1989). Some sparse, emerging evidence suggests that sonification as feedback may be immune to this effect (Mononen et al., 2003; Ronsse et al., 2011), and the current experiment was designed to investigate this question directly. Here, performance does not substantially decline after the removal of feedback according to any of the main measures, which might be taken as evidence that there is no guidance effect. However, sonification (either kind) did not produce a performance advantage relative to the Control condition during practice. It is equally likely therefore that intergroup performance convergence during practice (from a starting sonification disadvantage) represents participants in the sonification conditions learning to ignore the sonic feedback. It was intended that participants in the two sonification conditions could use extra sound information to guide movements more effectively than would be possible in their absence, but a lack of observed performance enhancement shows that this did not happen as intended. This leaves the central question of the guidance effect unanswered.

4.5.3 Unintended task constraints

This general pattern of results - either no advantage of sonification, or sonification as a hindrance makes some sense in light of the further analysis conducted on positional error of the initial movement. Motor task complexity can be defined in part by the constraints and requirements imposed by the testing environment (Kovacs et al., 2010; Wulf & Shea, 2002; Wulf, Shea, & Lewthwaite, 2010). Performance measurement during the acquisition stage (blocks 1-10) was done under differential conditions of auditory feedback. It was intended that the availability of extra information would be assistive, but instead it may have imposed additional, unhelpful constraints on movement. When moving to a target in the control condition, a participant can release the currently-controlled shape roughly on top of the target zone marker - this constitutes a successful move. In both the Full and Arrival Sonification conditions however, a successful move is signalled by the sound of a drum beat, which is triggered only when the measured position of a shape enters a zone defined around the marker. In effect, this mapping imposes a higher standard of positional accuracy upon shape movements in the sonification conditions than in control, as shapes need to be placed somewhat accurately in order to activate affirmative feedback. While positional accuracy was not a stated requirement of the task as communicated to participants, reproducing the demo animation (movement and sound components) was.

Further analysis revealed a between groups difference on the measure of positional accuracy on the first movement of each trial. At block 1, there was already lower positional error observable in the Full and Arrival Sonification conditions relative to Control, although pairwise comparisons did not reach statistical significance. By the end of practice this same pattern was still evident, with mean differences of similar magnitude, which now all reached significance. This indicates that sonification of corner arrivals in the two sonification conditions did impose an extra constraint on motor performance, requiring greater positional accuracy of arrivals. To speculate, a target-accuracy-focussed movement strategy may explain much of the poorer performance observed for the primary measures (global temporal error and cylinder harmonicity especially). Indeed, Bootsma et al. (2004) explain that when performing speeded movements towards targets, smaller targets tend to skew trajectory velocity profiles away from harmonic motion:

"*As the target aimed for becomes smaller, the duration of the deceleration phase (from peak velocity to end) lengthens, giving rise to an increasing asymmetry in the velocity profile that becomes elongated to the right*" (p. 814).

It can be said that the presence of beat-based arrival sonification (in both sonification conditions) represents the imposition of smaller targets. It may thus be possible to directly explain the acquisition-phase performance deficit in cylinder harmonicity in the sonification conditions.

Interestingly, the greater positional accuracy exhibited by participants in the sonification conditions during the acquisition stage was preserved into retention. At a test conducted five minutes post-practice, administered without auditory feedback, participants in both sonification conditions continued to move shapes to target zones with significantly lower positional error than participants in the Control condition. Despite removal of the constraint of sound, participants continued to perform the task in the style they had learned. A week later, the same pattern remained visible. Positional error was significantly lower for participants who had previously practiced with Arrival Sonification than control. A mean difference between Control and Full Sonification was present, but the difference was no longer statistically significant. Had positional accuracy been an explicit requirement of the current motor task, these results might be taken as evidence of superior motor skill learning with sonification, and evidence against a sonic guidance effect. However, positional accuracy was only part of task in the form of an unstated, implicit requirement of performance, and only then in the two sonification conditions. Despite these qualifiers, this finding is encouraging for the use of sonification as augmented feedback. Based on these results, it should be possible to encourage particular patterns of movement by making sonic feedback contingent upon achievement of particular task goals. In this case, the contingent sonic information and the movement pattern it implicitly encouraged were incidental to the stated aims of the task, but in another task, this arrangement could be harnessed to work to the advantage of the learner, and performance gains could be resistant to the removal of feedback.

4.6 Extended discussion

In retrospect, there are clear problems in: 1) the design of the motor task, and 2) the design of sound. A full accounting of both and how they intersect in participant experience should lay out a clear path forward for further research on sonification and motor skill learning. This discussion is partly a reflective exercise on the origins of the current experiment and begins with the motor task itself.

4.6.1 Task conceptualisation and mapping decisions

The development of the task used in the current experiment started from a set of unquestioned theoretical assumptions derived from the literature on augmented feedback. Very generally speaking, the classical research on knowledge-of-results and knowledge-of-performance feedback treats motor skill learning as the establishment of a motor program in memory (Adams, 1971, 1987; Salmoni, Schmidt, & Walter, 1984; for further, more critical discussion on this literature and its influence on modern research, refer to section 2.6.3 in the current thesis). This conceptualisation treats learning as fundamentally a top-down process, incorporating a central controller module which gathers sensory information and processes it internally to inform the construction of a motor plan. In this model, the motor plan is obviously of primary importance - therefore, coaching and feedback interventions should be geared towards its construction and maintenance. Although this literature is now very old, talk of 'internal models' and 'representations of the task' is still common in the contemporary literature on motor skill learning, including related work on movement sonification (see for examples: Effenberg, Fehse, Schmitz, Krueger, & Mechling, 2016; Kagerer & Contreras-Vidal, 2009; Oscari, Secoli, Avanzini, Rosati, & Reinkensmeyer, 2012). Following this, the current task was conceptualised as a *set of movements to be learned*, i.e. a sequence to be established in memory. Learning the movements was set as the goal, and performance metrics to measure achievement of movement goals were devised.

A problem with this theoretical approach to understanding motor skill learning is that it places a heavy focus on learning or acquiring actions, with only secondary consideration given to participant experience. An alternative conceptualisation (established in Chapter 2) states that skilful patterns of movement are enacted by a perceiving and acting agent. The learner must pick up information from the environment/task through interaction, and learn to recognise informational variables which could be controlled to achieve the intended task goals (Adolph & Kretch, 2015; E. J. Gibson, 1969). A manifestation of this theoretical divergence between learning movements and learning to perceive emerged in the current experiment, in which the chosen movement metrics did not adequately capture sonification's effects on the perception-action system. Sonification encouraged participants to move more precisely and attend more closely to accurate shape placement by constraining the conditions in which completion of that subtask could be perceived. This could perhaps be considered a constraint in line with the notion of a responsive perceptual-motor workspace (Newell et al., 1991). The finding of lower positional error in the sonification conditions indicates that in this task, the links between the primary performance metrics and participant intention were not as strong as they should have been. It is possible to further speculate that information specifying positional accuracy was more salient in participant experience (due to sonification) than information for achieving the stated task goals.

Thinking of the task as a top-down acquisition of movement patterns led to neglect of *how* participants would control their own movement performance, i.e. what was being perceived, and what (in terms of patterns of information) participants were trying to produce. Furthermore, the design of the task allowed for the use of many different perception-action strategies, and for attention to shift between many simultaneously competing subtasks (smoothness, timing, synchronicity, switching hands, directions, shapes, etc.). Given the multitude of subtasks, a strategy of 'sonify everything' was taken. In hindsight, an analysis of the task should have been performed, identifying the main features of the task from the learner's perspective (e.g. Wilson & Golonka, 2013). The informational variables controlled by novices and experts could perhaps have been identified and sonification deployed to aid pickup of information related to relevant features of the task.

Focussing more explicitly on the sound design in the current experiment reveals three additional problems. Firstly, many participants reported that they did not enjoy the sounds produced by the system - especially in the Full Sonification condition. Quite simply, there was "too much going on" - as reported by several participants. The sound of the system in motion was reportedly cacophonous and therefore demotivating to use for long periods of time. Participants were particularly tired of the sounds by the end of the second hour-long practice session. One of the major potential advantages of sonification over more traditional forms of graphical or numerical feedback lies in the notion that it might be intrinsically motivating to use, like an enjoyable session with a musical instrument. Learners could enjoy learning a new skill if movement produced interesting or pleasant sounds, an effect which could enhance motivation and therefore, learning (Wulf et al., 2010). According to reports from participants, this enhancement of motivation was not achieved with the current system.

The second sound-design problem in the current experiment is that the cacophonous nature of the system in motion could have made it difficult for participants to pick up information relevant to their currently-attended subtask. For example, it may have been difficult to audibly discern the trajectory of the slowly-moving cylinder shape over the sound of the prism and cuboid trajectories. The velocity of cylinder movement was perceivable using the sonic information provided through its movement, as movement up or down the workspace > 1.5 cm triggered a burst of filtered noise. The faster the movement, the greater the 'density' of the sound (a mapping metaphor inspired by movement over corrugated metal). The prism and cuboid trajectories were sonified when they need not have been: smoothness of either's trajectory was not a requirement of the task. Instead, sonification should have been restricted to only highlight information strictly relevant to the achievement of task goals. With only the cylinder shape sonified, perhaps its velocity could have been more easily perceived, and manipulated in line with the demo.

The third problem with sound design is applicable to the task as a whole. Deliberate efforts were made to utilise recommendations for action-sound mappings already established by the Auditory Display community and compatible action-sound pairings reported in previously-published work (Hermann et al., 2011; Rusconi et al., 2006). Sound mappings were designed for individual actions in such a way as to metaphorically evoke congruent 'real-world' actions (e.g. the directional pitch-mapping inspired by psychoacoustic functions derived from moving noisy objects, the metaphor of corrugated metal, a drum-beat to denote a 'hit'). Similarly, the three shapes were sonified using perceptually discrete sound types, in order that the activity of each might be independently perceivable. This is a well-accepted mapping strategy in the Auditory Display community when multiple streams of information are concurrently displayed (Flowers, 2005). Each stream of values is given a distinct auditory 'identity' in the hope that they will not perceptually overlap and corrupt listener understanding of the message. The use of this mapping strategy in movement sonification is commensurate with the 'top-down' approach to motor skill learning described earlier, in which perception is conceptualised as an internal process, separate from physical action. As described in section 3.2.3, a similar style of thinking about a 'disembodied' form of auditory perception remains the dominant form of discourse in Auditory Display (Roddy & Furlong, 2014; Worrall, 2010). Transposing the same thinking into the design of the current task and its sonification likely led to several of the problems explicated here. A conception of auditory perception as the passive reception of sensory information (to be processed later) led to the mapping decision to 'stream' the three shapes separately. The idea of motor skill learning as the construction of a symbolic motor program led to a series of mapping decisions intended to highlight the movements of the task in and of themselves, rather than the information that might be necessary to control movement in an 'online' fashion. Furthermore, there was no attempt made to use sonification to tie the whole task together as a meaningful auditory Gestalt (e.g. a melody, or recognisable auditory structure). If such an overall structure had been achieved, participants might have been able to better perceive the quality of their whole performance and perhaps achieved better performance with sound than without.

All of the points discussed above are reflective of misguided task-sound mapping decisions - related to the form of the task, and the sound design. In the flow of user experience, these were inseparable, and generally made for a difficult learning experience.

4.6.2 Learning bimanual coordination tasks

An important question which was not adequately addressed by the current experiment is: what informational variables were participants trying to control? Results from the positional error measure indicate that there was a disconnect between the main metrics of performance and participant experience of the task, i.e. participants in the sonification conditions perceived unexpected features of the task to be important, and moved accordingly. With a completely bespoke motor task, with so many moving parts, it is difficult to intuit how it might best be sonified to enhance performance - the most useful information to guide performance is unknown. There are however some tasks already well-enough researched in which it is possible to be reasonably sure. Bimanual coordination tasks have been touched on briefly prior to the current discussion; Ronsse et al. (2011) - one of the very few controlled motor learning experiments to contrast sonification as feedback against visual feedback - employed a bimanual 90° out-of-phase coordination task. Bimanual coordination tasks are often taken as a small-scale model of the perception-action system, and usually require a learner to maintain a complex timing or phase relationship between the hands (Shea, Buchanan, & Kennedy, 2016; Summers, Rosenbaum, Burns, & Ford, 1993). The relatively reduced nature of these tasks makes them an ideal vehicle for the exploration of motor control and learning in the lab. Coordination patters other than 0° (in-phase, or parallel movement) and 180° (anti-phase, or symmetrical movement) cannot be stably maintained by human subjects without extensive practice. These tasks are very difficult to learn, and augmented feedback is usually required (Kovacs, Buchanan, & Shea, 2009). In the past, this difficulty has been characterised as neuro-muscular in origin, i.e. due to limits of central processing for activations of different muscle groups (Kelso, 1984), however it has since been shown that learning how to perform well in bimanual coordination tasks is to a large extent a process of education of attention to perceptual information (Mechsner, Kerzel, Knoblich, & Prinz, 2001).

This insight was gained through an experiment in which learners' means of visually monitoring their hand movements was altered (via a geared mechanism) so as to appear as though they were producing a much simpler inter-manual phase relationship than they really were. In this way, a complicated physical action (circular oscillations at 4:3 relative frequency) produced an easy-to-perceive outcome (symmetrical 1:1 oscillations), and performance of the underlying physical task was enhanced (Mechsner et al., 2001). Mechsner et al. also report that a control group who attempted to learn the same 4:3 frequency task without visual transformation were completely unable to do so. This experiment showed that difficult bimanual coordination patterns are difficult to produce not because the brain is unable to produce the required motor commands, but because the correct pattern of motor coordination is difficult to tune into perceptually. When the means to perceive correct motor performance is simplified, the task itself becomes much easier. Further experiments have generally supported this view, showing that very difficult bimanual tasks can be learned with the right kind of transformed perceptual feedback (Kovacs et al., 2009; Kovacs & Shea, 2011; Wang et al., 2013). The perceptual feedback usually provided on these tasks is the Lissajous figure (for a full description of this system, see the introduction section of the current chapter, see also Figure 4.8 below). Task performance and speed of acquisition is much better when this feedback display is available. However performance invariably declines when Lissajous feedback is withdrawn; a classic demonstration of the guidance effect (Kovacs & Shea, 2011; Maslovat et al., 2009). This indicates that learners were controlling action exclusively using the visual information provided by the display when it was available, rather than intrinsic proprioceptive or visual-kinematic information from the limbs, which would be essential for transfer outside of feedback conditions. When the primary source of action-coupled information relevant to achievement of task goals is expressed via a feedback display, the display *is* the task, as far as the learner's experience is concerned.

Performance on bimanual coordination tasks improves even further when vision of the limbs is denied, and the display becomes the whole of the task. Kovacs et al. (2010) placed a screen over participants' hands, forcing participants to visually attend only to the display. It was found that they could learn to stably produce coordination patterns which had previously been assumed impossible (e.g. 5:3 frequency oscillations), with very little practice. This effect can be interpreted as the removal of a perceptual conflict: when limbs are visible, participants are forced to reconcile the oscillation of two limbs and the single swirling line of the display. The task is difficult because of competition for attention between two perceptually orthogonal systems. When one is removed (the hands are occluded), attentional resources are freed up and the learner can interact with a single system much more effectively than before. The physical movement of the limbs and the relative phase pattern between them is no longer a meaningful part of the task, therefore participants do not learn what good performance looks like in terms of hands moving relative to each other - they learn what the display looks like, and how to move so as to produce the target form on-screen.

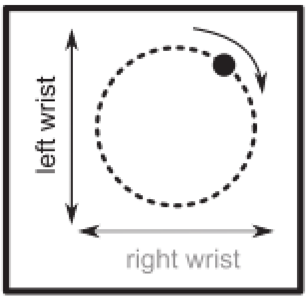

Figure 4.8: Lissajous figures plot displacement of one oscillator against another. This example schematic comes from Ronsse et al. (2011), in which 90° out-of-phase movement traced a circle.

4.6.3 Augmented feedback without transformation

There is some evidence which shows that the difficult underlying coordination patterns can be learned without transformation of the basic kinematics. Wilson, Snapp-Childs and Bingham, (2010) used a perceptual discrimination paradigm to train participants to visually recognise 90° out-of-phase oscillations on-screen before attempting to produce the same in a motor task. When these participants did attempt the motor task, they immediately showed a significant motor advantage relative to a control condition, in which participants had not undergone perceptual training. Furthermore, this advantage persisted into a later retention test. In a similar study, Wilson, Snapp-Childs, Coats, et al. (2010) trained participants to perform the same task, with non-transformed visual feedback indicating good performance. Participants practiced the task with full vision of the oscillators. When relative phase was within a predefined range close to 90° (a range which scaled with practice), the colour of a participant-controlled oscillator changed to green. This served as a signal that the task was being performed correctly, but did not disrupt the underlying perception-action dynamics of the task. Learning to recognise the information specifying when the task was being performed correctly allowed participants to improve at the task much more quickly than a control condition, who did not improve much at all over several practice sessions. Furthermore, when coloured feedback was withdrawn from participants in a later retention test, good performance persisted in that condition. This result again serves to highlight the strongly perceptual nature of motor control in bimanual coordination tasks, and shows that the performance-enhancing effects of feedback need not rely on transformation of task-intrinsic kinematic information.

The potency of the Lissajous figure as augmented feedback for bimanual coordination lies in the fact that it allows perception of a complex task as a single unified pattern. Participants can tell when they are performing well, because they are able to easily recognise the pattern for good performance, and how their own, current pattern relates to the target. With the Lissajous figure, very complex coordination patterns can be learned within minutes

(D. M. Kennedy et al., 2016; Kovacs et al., 2010). The problem (if lasting learning is desired) is that the intrinsic perception-action dynamics of the task are being altered, leading to a guidance effect. Wilson et al. show that the underlying task can be learned with non-disruptive feedback signalling when performance is within an acceptable range, but several sessions of practice are still required for learning. However there are yet other ways to manipulate the task situation in order to enhance information pickup and motor skill learning, while keeping the perception-action dynamics intact. Franz et al. (2001) designed a bimanual task which required participants to trace a semicircle with the index finger of each hand. The orientation of the required semicircles was manipulated in a way such that the motion of both fingers could be perceived either as two independent movements, or as a 'single task', i.e. a single unified Gestalt pattern (a whole circle). They found that when both hands moved together to produce a unified circle, between-hands interference was greatly reduced. They interpreted this finding as evidence that a bimanual dual-task can become, in learner experience, a single task - with the right kind of physical manipulation of the task environment. In this case, a circle represented a much more familiar and readily reproducible form than the alternative despite both versions of the task placing similar demands on muscular activation, i.e. a **()** task was easier than a **)(** task. In a more recent study, Franz and McCormick (2010) tested the efficacy of manipulations to the language used to cue participant action in a speeded bimanual reaching task with asymmetrically-positioned targets. Participants were told to either move 'both' hands (conceptualising the task in a unified fashion), or one hand 'and' the other (separate fashion). Although the movements required from participants were completely identical in both conditions, between-hands interference was almost completely eliminated in the unified condition, while still present in the separate condition. Although Franz and McCormick frame their explanation as 'conceptual', i.e. with reference to centrally-stored representations and concepts, these results are also commensurate with a purely perception-action framework. The language used for cueing can alter participants' relationship to the task such that they ready themselves to pick up and act so as to produce a specific kind of information (a more familiar, unified, Gestalt pattern), rather than a less familiar (i.e. less-learned) alternative. Participants may be outright told what kind of situation they are in, and constrain perception and action behaviour accordingly (see section 2.5). The task, including its performance and achievement of its goals can be perceptually unified without altering the physical layout of the task environment.

4.6.4 Concluding thoughts on participant experience

James Gibson, the Gestalt Psychologists and the Phenomenologists of the early 20th Century recognised that perception and action are unitary - and together, primary in experience (Dreyfus, 1996; J. J. Gibson, 1950; Humphrey, 1924; Kaufer & Chemero, 2015). Gibson built his Ecological approach to Psychology upon this basic tenet. When the information available for use in a motor task is altered, the task itself is changed in learner experience into something else - perhaps something more familiar, or easier to deal with. Franz et al. show that the same effect can be achieved by altering the learner's relation to the task, encouraging the adoption of a particular kind of readiness to interact with the world in a familiar way. Very often however, even within schools of thought which accept the unification of perception and action into a single framework, explicit talk of performer experience is relegated to the background in favour of things which can be more positively declared, like available information for phase relationships, the physical layout of the task, extant skills, or conceptual representations. The way the world (or, a task) is conceived in experience is intimately entwined with many of these explanations (Dreyfus, 1996). It may be possible to harness these ideas in the design of sonification for motor skill learning, such that learners' first-person experience of a motor task is structured by extant listening skill in the domain of music. The next three empirical chapters will test these ideas in a custom bimanual coordination paradigm and develop them further.

Chapter 5

Transposing Musical Skill: Sonification of movement as concurrent augmented feedback enhances learning in a bimanual task.

A version of this chapter is peer-reviewed and published under the same title: Dyer, J., Stapleton, P. & Rodger, M. *Psychological Research* (2016). doi:10.1007/s00426-016-0775-0

5.1 Abstract

Concurrent augmented feedback provided during acquisition can enhance performance of novel tasks. The 'guidance hypothesis' predicts that feedback provision leads to dependence and poor performance in its absence. However, appropriately-structured feedback information provided through sound ('sonification') may not be subject to this effect. This is tested directly using a rhythmic bimanual shape-tracing task in which participants learned to move at a 4:3 timing ratio. Sonification of movement and task demonstration was compared to two other learning conditions: 1) sonification of task demonstration alone and 2) practice with continuous pink noise (control). Sonification of movement emerged as the most effective form of practice, reaching significantly lower error scores than control. Sonification of solely the demonstration, which was expected to benefit participants by perceptually unifying task requirements, did not lead to better performance than control. Good performance was maintained by participants in the sonification condition in an immediate retention test without feedback, indicating that the use of this feedback can overcome the guidance effect. On a 24-hour retention test, performance had declined and was equal between groups. In the discussion, it is argued that this and similar findings in the

feedback literature are best explained by an ecological approach to motor skill learning which places available perceptual information at the highest level of importance.

5.2 Introduction

5.2.1 Movement sonification and the guidance hypothesis in motor skill learning

Concurrent augmented feedback is extra sensory feedback about movement which is presented live, alongside and during motor performance. It has been used successfully to enhance acquisition and learning in a wide range of motor tasks (Sigrist et al., 2013a). However learners typically become dependent on augmented information and performance declines when it is withdrawn (J. H. Park et al., 2000; Schmidt, 1991; Schmidt & Wulf, 1997; Sigrist et al., 2013b; Vander Linden et al., 1993). The high level of performance seen in the presence of concurrent feedback rarely persists into no-feedback retention tests, which constitute a truer test of learning (Salmoni et al., 1984). This may happen when learners come to rely too heavily on the augmented information provided by concurrent feedback, and ignore task-intrinsic sources of sensory feedback, an effect known as the ‘guidance hypothesis’ (Adams, 1971). Once augmented feedback is removed, the learner must rely on comparatively unfamiliar sources of intrinsic feedback (e.g. proprioception) and performance declines as a result of impaired performance-monitoring ability (Anderson et al., 2005). Intrinsic sources of sensory feedback may be unattended when augmented feedback is available for two possible reasons. 1) The feedback display may simply distract attention from otherwise available intrinsic information, or 2) it may provide performance information which is much easier to use than intrinsic sources[23].

Emerging evidence suggests however, that the guidance hypothesis is not a general principle of feedback as had previously been assumed (Danna et al., 2014; Mononen, Viitasalo, Konttinen, & Era, 2003; van Vugt & Tillmann, 2015; for a review, see Chapter 2). Experiments using concurrent feedback in the auditory modality have shown that speed of

[23] The latter is likely the case for certain kinds of transformed visual feedback, which come to stand in for very difficult-to-use intrinsic sources (e.g. Kennedy, Wang, Panzer, & Shea, 2016; Mechsner et al., 2001).

acquisition can be enhanced using sound without impairing performance on subsequent no-feedback retention tests (D. M. Kennedy, Boyle, & Shea, 2013; Ronsse et al., 2011).

For example: Mononen, Viitasalo, Konttinen, and Era (2003) sonified one-dimensional aiming error in rifle training by mapping positional error of the gun barrel to sonic pitch. Their participants therefore had access to an additional layer of performance-relevant information through sound and performance was improved as a result. Unlike concurrent feedback experiments in the visual modality, no decline in performance was observed following the removal of augmented feedback. The enhancement effect of feedback was maintained on no-feedback retention tests, even several days later.

Ronsse et al. (2011) tell a similar story and provide a rare example of visual and auditory concurrent augmented feedback contrasted on the same experimental task (90-degree out-of-phase bimanual flexion/extension). Concurrent visual feedback was provided in the form of a Lissajous figure and auditory feedback via sonification of changes in wrist direction, which results in a 'galloping rhythm' when movements are performed accurately. They found that although visual feedback allowed learners to reach optimal performance more quickly than auditory feedback, this high level of performance was maintained only by the auditory group in no-feedback retention. A typical guidance effect was found following the removal of visual feedback, but not auditory feedback. Heitger et al. (2012) replicated the behavioural findings of Ronsse et al. using the same bimanual task.

These findings represent a slight challenge to traditional interpretations of the guidance effect, which assume that feedback presented 100% of the time during acquisition will lead to decline when it is withdrawn because intrinsic proprioceptive feedback has been attentionally neglected (Anderson et al., 2005; Sigrist et al., 2013a). However these results make a lot of sense from a perception-action perspective.

5.2.2 A perception-action perspective on the guidance effect in bimanual tasks

Motor learning in bimanual coordination tasks is clearly perceptually-based[24] (Franz et al., 2001; Mechsner et al., 2001; Wilson, Snapp-Childs, Coats, et al., 2010). Bimanual coordination performance is so difficult to perceive intrinsically that learner attention is occupied entirely by controlling the feedback display; this is by far the most valuable information that the environment offers in the context of the task – and guidance effects are the norm (Kovacs et al., 2009; Kovacs & Shea, 2011). In this situation, the learner does not actually learn to produce the bimanual task; he/she learns how to manipulate the display. It is very difficult to perceive useful information about bimanual coordination from the limbs themselves, and in fact any such information may actually conflict with the Lissajous information (Kovacs et al., 2010).

The guidance effect then comes as no surprise. In the case of visual feedback, the display *is* the task. This fact is not of great concern if one's goal is to push the limits of perceptual control of action (e.g. Kovacs et al., 2010), but it is a real problem if the aim is to produce learning which transfers outside the lab. If the only way (or, the most effective way) for the learner to perceive their performance is through an augmented feedback display, then he/she will not be able to perform the task in its absence. In the next section, movement sonification will be examined from the same perspective.

5.2.3 Noisy events, perceptual unification and sonification

Sonification is (or rather, can be) more than just another method for abstract display of symbolic movement data (Roddy & Furlong, 2014). There are distinct perceptual and phenomenological qualities of sound perception which may make it a more appropriate modality for meaningful concurrent feedback than a visual display (Dyer, Stapleton, & Rodger, 2015). These qualities can explain sonification's potential immunity to the guidance effect.

[24] This is likely also the case for motor skill learning in general. Bimanual coordination is not a special kind of learning except that the effects of manipulation of perceptual information are much more profound than in most other tasks.

Sound is intrinsically linked to movement in experience (Leman, 2008; Repp, 1993; Sievers et al., 2013). In everyday life, sounds automatically become part of multimodal event perception (Gaver, 1993). Thanks to our extensive interactive experience with a noisy environment, we can perceive a surprising amount of action-relevant information from an auditory event (Giordano & McAdams, 2006; Houben et al., 2004; van Dinther & Patterson, 2006; Young, Rodger, & Craig, 2013). In the case of sounds produced by action, fMRI studies during passive listening have recorded neural activations similar to those observed during previous action performance (Kohler et al., 2002; Lahav et al., 2007). Behavioural effects are especially strong for extensively-practiced noisy actions, for example instrumental performance (Taylor & Witt, 2015). Additionally, specific actions can even be identified from their sonified velocity profile alone (Vinken et al., 2013). Summarised, sound and movement are ecologically coupled. Sound is inherently meaningful to the moving individual, and if it were employed as concurrent augmented feedback in a motor skill learning study, the link between participant movement and feedback could potentially be much tighter, and feedback less of an abstraction. In other words, sound as feedback can be more coupled to fundamental task kinematics than a visual display. The use of sound can perhaps more explicitly include the body in the perception-action loop.

As shown by Ronsse et al. (2011) and Kennedy, Boyle and Shea (2013), auditory models/demonstrations of bimanual task performance along with sonification as feedback are effective for training complex coordination tasks. Making perceptual information about bimanual task performance more salient or perceivable leads to reduced variability in associated action (Wilson, Collins, & Bingham, 2005). This seems to be a general perceptual effect which also applies to sound information and unimanual tasks. van Vugt and Tillmann (2015) found that accurate sonic feedback improved tapping accuracy in a learned motor task to a greater degree than jittered feedback. Interestingly, improved performance in the sonification group persisted into no-feedback retention and transfer tests. The temporal resolution of the auditory system is known to be much finer than that of the somatosensory system (Hirsh & Watson, 1996; Tinazzi et al., 2002), so one would expect more accurate temporal perception of any event paired with sound. Following a perception-action approach to motor skill learning, and assuming that perception never happens in isolation from action (E. J. Gibson, 1969), it stands to reason that enhanced perceptual acuity for action's consequences (i.e. feedback) will necessarily result in better control of action.

Ronsse et al. show that, although slightly slower, sonification is *as effective* for teaching a novel coordination pattern as the more commonly-used Lissajous figure. Lissajous feedback works through perceptual unification, a transformation wherein a difficult bimanual task is consolidated and abstracted to create a new, more coherent and unitary percept (for the effect of perceptual unification on other bimanual tasks without Lissajous feedback, see Franz et al., 2001; Mechsner et al., 2001). Unification makes relevant perceptual information about the higher-order variable of relative phase/timing ratio more available, which allows effective and stable action production. It is likely that a demonstration through sound functionally does the same thing; *it consolidates a dual-task into a rhythm*, which can be perceived and reproduced as a single action.

The potential advantage of sonification over Lissajous *as concurrent feedback* lies in the degree of abstraction, or transformation. As argued earlier, and presupposing good sound design[25], sonification of bimanual coordination does not entail the same degree of transformation as does feedback displayed as a Lissajous figure, the Gestalt form of which differs substantially from the underlying kinematics of bimanual coordination. By contrast, sonification is layered on top of and can be used to emphasise relevant task kinematics. This style of feedback can allow direct perception of phase relationship or timing ratio without subsuming the main motor task, as recommended by Wilson, Snapp-Childs, Coats and Bingham (2010). Information related to the higher-order relationship between the hands *is present* in task-intrinsic proprioceptive feedback; it should be possible to use sonification to train participants to perceive it directly – eliminating the guidance effect of concurrent feedback.

---

[25] Sound design is often given only cursory attention in perceptual-motor learning studies dealing with sonification. A case for its importance will be presented in the discussion.

The aim of the present paper is twofold. Firstly, it aims to further scientific understanding of the guidance effect of concurrent feedback, specifically how it relates to sonification. Secondly, it aims to separate the effects of perceptual unification from feedback in order to test whether unification of the task goals (through adding sound to the demonstration) is sufficient to enhance learning, or whether there is a distinct advantage of sonification as concurrent feedback. At this point, it is not yet clear whether the effects of sound on learning in Kennedy et al. (2013) are due to either perceptual unification through a sonic demonstration, or concurrent movement sonification. Performance in bimanual coordination is improved by perceptual unification alone (Franz & McCormick, 2010; Franz et al., 2001), and it will be important to establish this difference going forward. After all, one need not provide online sonification of movement during practice at all if performance can be enhanced to the same degree by using a pre-recorded, sonified demonstration.

To this end a novel bimanual shape-tracing apparatus was designed to teach participants to produce a 4:3 rhythmic coordination pattern, a task previously shown to be difficult to learn (Summers et al., 1993).

It is hypothesised that the use of sonification as auditory feedback will not lead to a guidance effect relative to no-sound control. Like Lissajous feedback, sonification represents a method to perceptually unify a bimanual task; however it does not rely on a transformation and abstraction of the fundamental task kinematics. For this reason both enhanced performance of the sonification group during practice, and maintenance of this enhanced performance into retention-without-feedback are expected.

It is additionally hypothesised that performance in the condition in which the demonstration alone is sonified (hereafter referred to as the ‘sound-demo condition’) will benefit from the use of sound to perceptually unify the task demands, which will manifest as enhanced performance during practice and into retention relative to no-sound control.

Performance with the sound-demo alone and sonification as concurrent feedback will also be compared. Both conditions perceptually unify the task demands, however live sonification may confer a relative advantage in the acquisition stage by enhancing online temporal perception of performance. Improved perceptual acuity through sound should, in general, manifest as better performance (Fowler & Turvey, 1978), and the same is expected in this task, good performance in which is based at least partly on fine temporal control.

5.3 Methods

Participants

An opportunity sample of 45 right-handed participants (20 female; mean age = 24.3 years [S.D. = 5.9 years]) was recruited from a combination of undergraduate Psychology students, postgraduate researchers and staff at the university in which the experiment was conducted. Undergraduate students received partial course credit for their participation. Right-handedness was confirmed for all participants by administration of the Edinburgh Handedness Inventory (Oldfield, 1971). Handedness scores did not differ significantly between experimental groups ($F(2,42) = .335$, $p = .717$).

Participants were questioned about their musical experience after completion of the study to avoid experimenter bias. Almost half (21 of 45 participants) reported some experience playing musical instruments, in most cases not currently. Eight participants in the Sonification condition reported musical experience, only one of whom was active. The other 7 reported having ceased playing an average of 5.4 years ago. There were 6 musical participants in the Control condition, 4 active, the rest having ceased mean 5.5 years ago. The Sound-Demo condition contained 7 musical participants, 1 active, with the rest having ceased mean 3.5 years ago.

Informed consent was obtained from all individual participants included in the study. Ethical approval for this study was granted by the School of Psychology Ethics Board at Queen's University, Belfast.

Materials and Apparatus

Hardware: A bespoke wooden board (70cm x 30cm) was created for the purpose of this experiment (see Figure 5.1). Two 20cm x 20cm slots were cut into the top side of the board, into which were inserted a pair of wooden slabs. On each of the slabs was carved a regular polygon (a diamond on one and a triangle on the other) of equal path length (34cm). Shape grooves were rounded with 3mm depth (at the centre) and 12.5mm width. The board was placed on a desk at which participants were seated. Participant movement data were obtained using a Qualisys optical motion capture system capturing at 300Hz, which was triggered using an Arduino controller. Participants wore a pair of modified golfing gloves with reflective markers attached, allowing the movement of the hands and tip of the index finger to be tracked in 3D space.

Figure 5.1: Participants traced the index finger of both hands around the shapes simultaneously in an anticlockwise direction, starting from the top corner.

A 17-inch screen was used to display a demonstration animation corresponding to exemplary performance and a pair of Sennheiser headphones were worn by participants at all times. The experiment was administered by the experimenter using a desktop PC running *Qualisys Track Manager (QTM).*

Software: Data corresponding to participant movement in Cartesian space (x, y and z) were streamed in real-time from *QTM* to *Max/MSP 6.0* via the *Open Sound Control* (OSC) protocol. An exemplary demo animation and graphical display were programmed using *Processing.*

Sonification and terminal feedback: In this experiment, participants engaged in a series of discrete practice trials, following which, post-trial (terminal) feedback was provided. A 3x3cm (9cm^2 area) range was defined for each corner of the diamond and triangle shapes (i.e. a square, centred on each corner, boundaries extending 1.5cm bi-directionally in the x and y planes), based on the position of the index finger marker (x, y) when a participant's fingertip was positioned in the corners. A trigger was produced in *Max/MSP* by index finger arrival in any of these zones. An inter-trigger interval (time between corner arrivals) was thus calculated for the left and right hand. Each new right-hand interval was compared to the previous interval for the left hand to calculate a ratio (with the target right-to-left duration ratio of 3:4). These ratios were stored and displayed on a graph at the end of each practice trial as terminal feedback (See Figure 5.2).

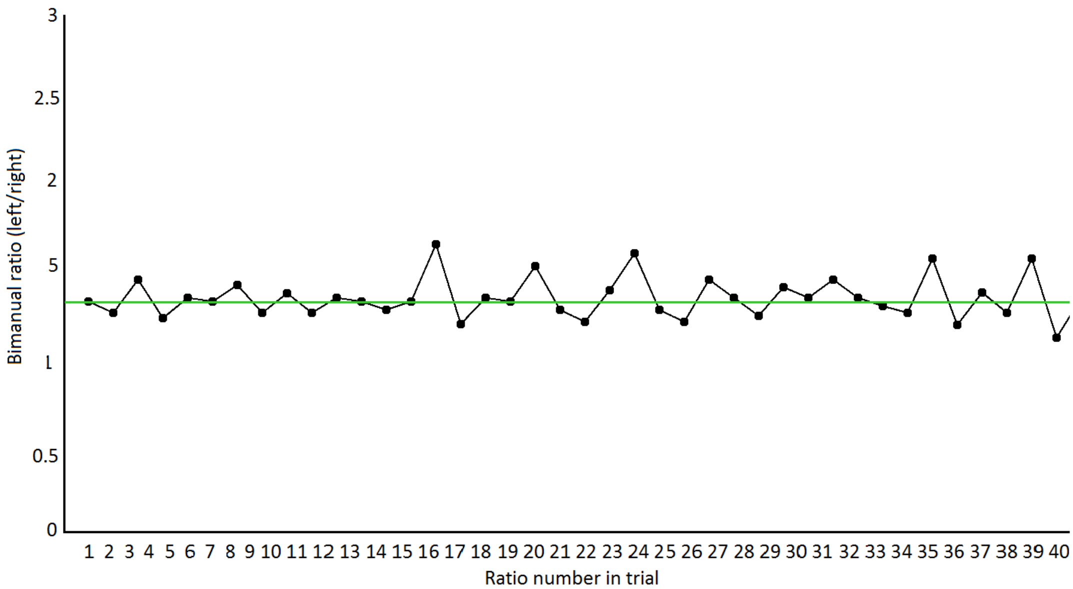

Figure 5.2: Intermanual ratio was continuously plotted on a graph which also showed the ideal 4:3 (1.33) ratio as the horizontal midline. The graph shown corresponds to relatively good performance (low error magnitude and variability). Axes labels were not visible to participants.

These same arrival triggers were used as the basis for concurrent sonification feedback. This model of sonification draws some inspiration from Ronsse et al. (2011), who sonified reversals in direction in a bimanual task; the endpoint of a movement trajectory was judged to be a salient perceptual event in both Ronsse et al. and the current experiment, and tightly-linked to the main goal of the task, i.e. intermanual movement frequency ratio. In the current experiment, each endpoint of a movement trajectory (i.e. arrival at a given shape corner) was represented by one of a set of notes in the key of C Major. Tones were generated in *Max/*MSP by combining a pure tone (with a given frequency corresponding to one of the notes in Fig 5.3) with a predefined envelope function which modulated loudness over time. Following a trigger which initiated the tone, loudness decayed roughly exponentially, reaching silence after 350ms[26]. The notes for the left and right hand were taken from separate but adjoining octaves, as a close pitch relationship has been shown to be conducive to auditory "stream" formation and perceptual integration (Bregman & Campbell, 1971; Flowers, 2005). Thus a short melody was played by correct performance of the task.

Figure 5.3: The left and right hand corner arrivals were sonified using synthesised tones not associated with any real-world instrument. The left (bottom) and right-hand (top) movements were unified into a single melody when the task was performed correctly.

Procedure

Participants were pseudorandomly allocated to one of three conditions: Control, Sound-Demo and Sonification (N = 15 each). Each of these conditions entailed different availability of sound to guide performance. For a graphical visualisation of the entire experimental procedure, please refer to Figure5.4.

Familiarisation

[26] A link to a video showing performance of the task and associated sonified feedback is available in Appendix A.

The experiment began with a short task-familiarisation phase in which participants in all three conditions were shown a soundless visual demo animation of correct task performance. The demo showed two shapes on-screen (corresponding to the wooden shapes in front of the participant). Individual corner zones of the animated shapes lit up in sequence, demonstrating the spatio-temporal characteristics of the required 4:3 bimanual coordination ratio (Hove & Keller, 2010). One full cycle of the demo lasted 3 seconds (a left inter-trigger-interval of one second, right 750ms). Three rotations were presented on each 'play' of the demo. Participants were played the demo twice during this familiarisation phase (comprising 6 rotations in total), then given approximately 15 seconds movement time, in which they attempted to reproduce the spatiotemporal characteristics of the movement seen in the demo. Participants in the sonification condition had their hand movements sonified during this time which served as familiarisation with the action-sound mapping; however no participants had access to an audible demo at this point.

Practice

The practice phase consisted of 14 discrete trials for all participants. Each trial began with a play of the demo (9 seconds), followed by a movement phase (26 seconds), and concluded with presentation of terminal feedback (see Figure5.2).

The Control condition saw a purely visual demo and listened to constant pink noise during its presentation. During the movement phase for the Control condition, no sonification was provided – only constant pink noise was heard. Pink noise was used (at low volume) during the movement phase to mask any naturally-occurring sounds from hand movement over the apparatus. Trials concluded with the graph presented as terminal feedback.

The Sound-Demo condition saw a visual-acoustic demo at commencement of each practice trial, in which corner arrivals were sonified using the tones shown in Figure5.3, without pink noise. During the movement phase, participants heard constant pink noise. Trials concluded with the graph.

The Sonification condition saw the same visual-acoustic demo as the Sound-Demo condition at commencement of each practice trial, without pink noise. During the movement phase, arrivals of the index fingers at corner zones were sonified using the procedure described earlier and the notes in Figure5.3. Perfect performance of the 4:3 ratio would produce the same melody heard in the demo. No pink noise was heard during movement. Trials concluded with the graph.

Retention

After 14 practice trials, all participants were given a five-minute break before undergoing a 26-second retention test without any augmented feedback (i.e. no graph and no sonification – where applicable). No demo was played prior to this trial and no graph was shown afterwards. Participants in all three conditions heard pink noise during the movement phase. The retention test was repeated exactly on the following day.

Transfer

Lastly, a transfer test was administered to assess whether task learning would generalise to a differing degree based on the mode of learning. The application of learned motor skill to a different task context is generally taken as an indicator of robust learning (Soderstrom & Bjork, 2015). We tested transfer by switching the positions of the shapes to be traced. The task was essentially the same; 4:3 rhythmic coordination, only mirrored.

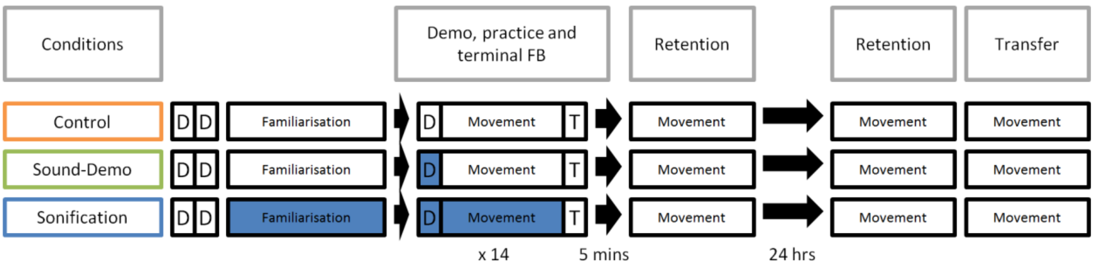

Figure 5.4: Experimental procedure. Boxes marked **D** represent a presentation of the demo animation. Boxes marked **T** represent terminal (graph) feedback. Blue/shaded boxes indicate the presence of sound at corner arrivals/sonification. All unshaded movement and demo sections occurring after familiarisation were paired with constant pink noise.

5.4 Results

5.4.1 Bimanual ratio of timing

Bimanual timing ratio was calculated continuously for each trial by comparing every right hand inter-trigger interval to the most recent interval for the left hand. This raw information was presented to participants as terminal feedback. For analysis, the difference between the values of these obtained ratios and the ideal (4:3) ratio was calculated, yielding a measure of absolute error over time. The mean of absolute ratio error served as a measure of performance for each trial, with a value of 0 indicating trial performance which perfectly matched the target ratio throughout.

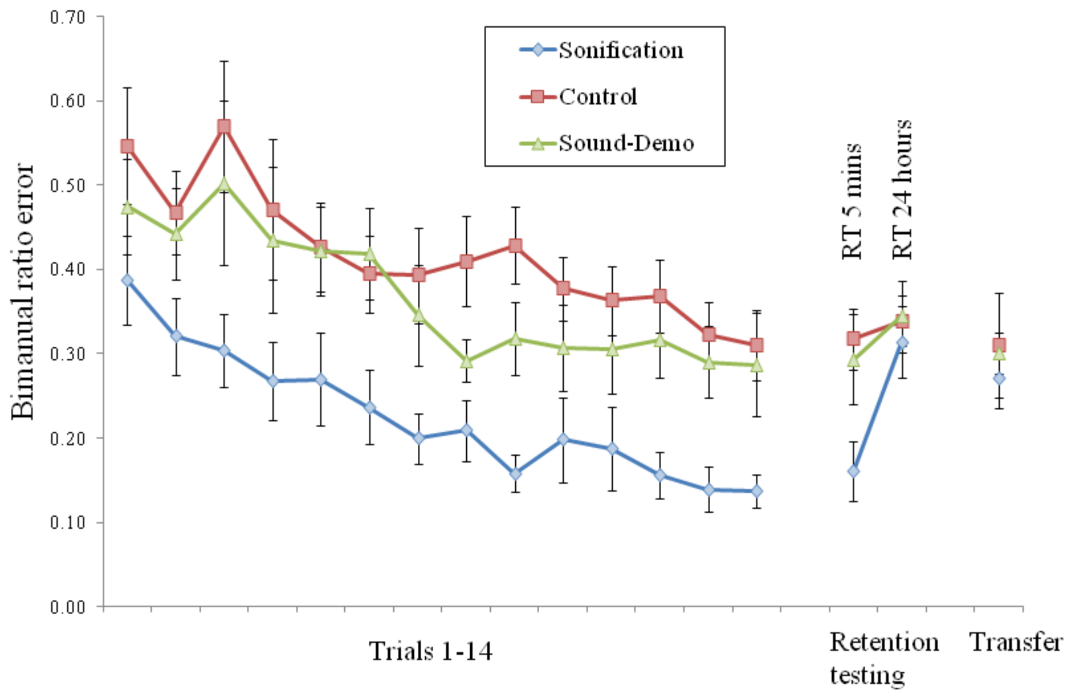

Figure 5.5: (Learning curves) Rates of average absolute ratio error for the three feedback groups during practice, retention and transfer. A score of 0 represents perfect performance. Feedback was provided on trials 1 – 14. Error bars are standard error.

A mixed ANOVA on acquisition data (trials 1 – 14) with condition as a between-groups factor and trial as a repeated measures factor revealed a significant main effect of condition: $F(2, 39) = 6.75$, $p = .003$, $\eta^2 = .137$ and trial: $F(5.098, 198.804) = 12.29$, $p < .001$,

$\eta^2 = .110$. No trial x group interaction was detected: $F(10.200, 0.298) = .423$, $p = .936$. Pairwise comparisons of inter-group score differences were performed at Trial 14 to test whether there was a significant benefit of sonification by the end of practice. Alpha was set at .016 (Bonferroni correction for three comparisons). The Sonification condition performed the task with significantly lower error than the Control condition ($p < .001$, Cohen's $d = 1.344$), but not the Sound-Demo condition ($p = .031$, Cohen's $d = .839$). No significant difference in scores was evident between the Sound-Demo and Control conditions ($p = .757$). Participants who learned with sonification were evidently better at the task on the final practice trial than their counterparts in the Control condition.

To identify differences in rates of learning, curves were iteratively fitted to learning curves as described in the previous experimental chapter (section 4.4.4) and slope values compared between groups. Holm-Bonferroni corrected *p* values are reported for this analysis. In the Sonification condition, the mean slope value was .590 (S.D. = .664). A one-sided *t*-test (against 0) revealed that slopes were significantly greater than 0: $t(14) = 3.033$, $p = .020$. In the Sound Demo condition, the mean slope value was .928 (S.D. = 2.974). A one-sided *t*-test revealed that slopes in this condition were not significantly greater than 0: $t(14) = 1.209$, $p = .369$. In the Control condition, the mean slope value was 1.103 (S.D. = 2.782). A one-sided *t*-test revealed that slopes in this condition were not significantly greater than 0: $t(14) = 1.536$, $p = .735$. ANOVA across conditions revealed no effect of condition on slope values: $F(2, 44) = .180$, $p = .836$.

One of the primary interests in the current experiment is in the presence (or absence) of a guidance effect after the removal of sonified augmented feedback. It is crucial to be able to tell whether the improved performance in the Sonification condition was dependent on the presence of feedback, and whether it deteriorated after it was removed. To this end, a test of statistical noninferiority was performed on Sonification group error scores in the 5-min retention test relative to Trial 14. This procedure is described in full by Walker and Nowacki (2011). In brief, if the 90% confidence interval (CI) of the difference scores (between trial

14 and 5-min retention) falls within a pre-set noninferiority interval, then noninferiority of retention performance can be inferred at the .05 level. The noninferiority interval is here set at .087, given that this is .5* the difference in mean scores between Sonification and Control conditions at trial 14. If the upper CI of the 5-min retention minus trial 14 difference scores falls below this value, then it can be said that performance did not deteriorate (positive values indicate performance worsening in this arrangement). This is a common procedure for noninferiority testing in clinical drug trials in which noninferiority of a new drug (relative to an old drug) is inferred based on whether the 90% CI of difference scores between a new drug and the old falls within an interval set by 0.5* the difference between the efficacy of the old drug and placebo (Walker & Nowacki, 2011, p. 194). The mean of the difference scores between Trial 14 and 5-min retention was 0.021, with a 90% CI of [-0.041, 0.062]. The upper CI is below the non-inferiority interval of .087, which means that performance was not inferior after sonification was removed. A p value for the noninferiority test can also be provided (as recommended by Walter and Nowacki) by performing a one-sided, one-sample t-test on difference scores relative to the equivalence interval, 0.087: $t(14) = -2.841$, $p = 0.013$).

On the second retention test, it is clear from Figure 5.5 that the advantage of sonification had evaporated and performance had declined. Testing for group differences at this point revealed no main effect of condition $F(2,42) = 4.15$, $p = .663$, $\eta^2 = .020$, indicating that between-group performance had equalised at this point. Performance was similar on the transfer test, where no main effect of condition was present $F(2,42) = 1.29$, $p = .287$, $\eta^2 = .054$.

5.5 Discussion

5.5.1 Benefits of sonification for performance in acquisition

By the end of acquisition, participants in the Sonification condition showed improved performance relative to Control, which indicates that concurrent sonic feedback was beneficial for acquisition (see Figure 5.5). In this experiment, sound was used for two task-relevant purposes: one, to allow participants to directly perceive the higher-order variable which constituted the main goal of the task: bimanual timing ratio. This was accomplished by attaching tones to corner activations in the demo (and practice for the Sonification condition), creating a global melodic pattern (Franz & McCormick, 2010). Two, to more precisely specify (temporally speaking) the micro-level structure of the pattern i.e. the required timing of individual corner arrivals (and produced timing, in the case of Sonification). It has been shown that the temporal-perceptual resolution of proprioception is much lower than that of audition (Hirsh & Watson, 1996; Tinazzi et al., 2002), and it was hoped that this could be augmented by exploiting sound to more clearly specify the temporal position of each corner-arrival. The performance data from the Sonification condition then conform to the hypotheses. Sonified participants had access to a both a unified percept of the required movement pattern and precise temporal specification of their performance, an arrangement which facilitated very fine-grained performance-monitoring and demo comparison. As can be seen in Figure 5.5, the Sonification group showed improved performance relative to both other groups throughout the acquisition phase.

Some differences in rate of learning were also observed between experimental groups. Analysis of the slope of curves fitted to performance data from individual participants in the Sonification condition showed that learning rates were significantly greater than 0. This was not found for corresponding curves fitted to participant data in the Sound Demo and Control conditions. The lack of a stronger difference here may be due to a limitation of the experimental design, which does not include a true pre-test under identical experimental conditions across groups (see Fig 5.4). Instead, the first trial for the Sonification condition included the presence of sound feedback, and the demo was immediately sonified in both the

Sonification and Sound-Demo conditions. It is therefore inappropriate to treat the first trial as a pre-test or baseline measure of performance. Although performance on the first trial was extremely variable between participants, it is possible that an immediate first-trial advantage for Sonification was in play. This could have caused the learning curves to appear slightly more parallel and similarly-shaped than they would have been had a true pre-test been conducted prior to trial 1.

Given the high informational value of sound in this context with regard to demonstration, the finding that there was no corresponding advantage evident in the Sound-Demo condition relative to Control by the end of acquisition was unexpected. D.M. Kennedy et al. (2013) found that practice with an auditory model led to lower error and variability than with a purely visual model, and the same had to some extent been expected here, despite the confounder of concurrent auditory feedback in Kennedy et al. Instead, highly similar performance in the Sound-Demo condition to Control at trial 14 was found, and similar rates of performance improvement from trial 1–14.

The factor which differentiates Sonification then, is the availability of *concurrent* auditory information. Participants in the Sonification group completed fourteen 26-second-long trials of a novel, semi-musical movement task, which seems to have been enough practice to learn the mapping between action and sound. A merging of perception and action occurs in musical instrument training (Drost, Rieger, Brass, Gunter, & Prinz, 2005), such that actions are perceived in terms of their musical outcomes – and it is likely that a similar merging occurred here, despite the comparatively brief timescale. The movements of the motor task were mapped to musical tones; this is a simple, tightly deterministic mapping not unlike that of a traditional musical instrument (for more detailed discussion on mapping strategies in digital instruments, see section 3.3.7). In summary, the working mapping enabled participants to use auditory information to better perceive and coordinate motor performance.

This may explain why the Sonification condition showed an advantage in performance relative to Control when the Sound-Demo condition did not. The relatively low temporal acuity of proprioception as a feedback modality may have been a limiting factor for performance in the Sound-Demo condition, whereas proprioceptive feedback was augmented with sound in the Sonification condition. As predicted by a perception-action approach to motor control (Fowler & Turvey, 1978; E. J. Gibson, 1969), enhanced perception of action's consequences led to improved control of action.

A specific benefit of Sonification relative to the Sound-Demo by the end of practice (on trial 14) may have been expected. Although the difference between groups at this point was in the expected direction (ratio error of .14 and .29 in Sonification and Sound-Demo respectively), a post-hoc t-test did not quite reach statistical significance ($p = .031$, $\alpha = .016$). This result can perhaps be attributed to relatively high performance variability in the Sound-Demo condition at this time (S.D. = .24, compared to .08 in the Sonification condition), making statistically significant mean differences between the Sound-Demo condition and others more difficult to detect. This may also indicate that the study was slightly underpowered, necessitating recruitment of a larger cohort of participants in future research of this kind.

5.5.2 The 'guidance effect' in early retention

A very similar pattern of results can be observed in the first no-feedback retention test as appeared on the final practice trial, when feedback had been available. Good performance by the sonification group was carried over into retention. Participants were able to overcome the guidance effect of concurrent feedback and maintain good performance without sonification. This result is in accordance with Ronsse et al. (2011) and Heitger et al. (2012), who also found no evidence of a guidance effect upon removal of auditory feedback in a bimanual task.

This finding suggests that participants had been trained to more accurately perceive the higher-order variable of bimanual timing ratio from their own intrinsic feedback, as expected. The production of a melody specified the required pattern, essentially making retention a mute musical recital. It has been reported that musicians experience sounds associated with practiced musical actions when performing the actions in isolation (Lotze, Scheler, Tan, Braun, & Birbaumer, 2003), and that this audio-motor coupling can be induced in amateurs with relatively little practice (Lahav et al., 2007). This lines up well with post-experiment reports from participants in the Sonification condition, almost all of whom stated that they imagined playing the melody during the first retention test. van Vugt and Tillmann (2015) found that sonification of finger tapping resulted in improved timing accuracy and lower tapping variability (a result predicted by the perception-action approach to motor control invoked earlier), however the benefit persisted even after the removal of sound. This implies that learned links between sound and movement may allow such experienced individuals to more accurately perceive temporal information in proprioceptive feedback, overcoming its intrinsic limitations. Thus it is possible that a coalition of benefits associated with sonification were in operation in the current experiment to produce this result.

This study then adds to the growing literature on sonification and its apparent immunity to the guidance effect (Heitger et al., 2012; Mononen et al., 2003; Ronsse et al., 2011; Sigrist et al., 2013a; van Vugt & Tillmann, 2015).

5.5.3 Exploiting the musicality of movement in retention

At the 24-hour retention test, no benefit of Sonification relative to Control or the Sound-Demo conditions was observed. ANOVA revealed no significant effect of condition at this point, as performance in the Sonification group roughly equalled that of the two others. Reports from sonified participants at the time of this test indicated that most could no longer remember what the melody was supposed to sound like, and were keenly aware that their performance had declined from the previous day, despite receiving no feedback of any kind. It is thus unsurprising that the same pattern of results is observed in the transfer test, which was conducted immediately following the 24-hour retention. Participants had lost the ability to perform the base task; therefore they were unable to apply their skill in a novel scenario.

This suggests that this 4:3 bimanual coordination pattern had effectively become a musical task. The ability of music to guide movement in a way which is aligned with more abstract task goals represents the fundamental (and currently underexploited) potential of sonification to train a wide range of otherwise non-musical skills. From a motor-performance perspective, accomplished musical instrument performance represents one of the most impressively complex and temporally-precise ways in which the human motor system can be deployed. This deployment is of course in service of a higher-order goal, the production of music; an accomplished performer is generally less concerned with the minutiae of motor control at the muscular level than the creation and maintenance of an overall Gestalt in the form of music. This is evidenced in the observation that the types of errors made by more advance-skilled musicians are those which are more likely to preserve the harmonic and temporal integrity of the musical whole (Drake & Palmer, 2000). Furthermore, we can be certain from the perception-action literature discussed here and in Chapter 2 that the precision of motor output evident in musical performance is afforded precisely *because of* the audio-motor link inherent in music. The recruitment of auditory perception in concert with a process of learning which enables an understanding of how one’s movement can alter sound, results in control of motor output which is unrivalled in most other domains of activity. The present experiment shows that potential exists for the exploitation of music in motor skill learning (through sonification), which in theory could be applied to many other skills that require precise control of movement, e.g. sport, or re-learning of basic skills in motor rehabilitation. If the latent musicality in skilled action can be brought out, movement sonification could see broad applicability in skill learning.

Further research should focus on ways to extend sonification's guidance-effect immunity in time; one could speculate for example that refreshing a learner's memory as to the exemplary sound profile might enable early retention-level performance to re-emerge, as perception of the sonic outcome of musical motor performance may enable amodal/holistic perception of the movement event which precipitated it (see section 2.7.3). Performance could thereby be enhanced without actually ever needing to re-expose participants to concurrent feedback.

5.5.4 Conclusion

The main finding in the reported experiment concerns the guidance effect of augmented feedback as it applies to sonification. It has been shown that, under the right conditions, concurrent sonification can overcome the assumed dependency on feedback. This was possible by treating the task as a musical one, which allowed participants to display some of the fine temporal and higher-level Gestalt control of movement commonly seen in musical instrument performance. Similarly to how accomplished piano players can produce reasonably accurate performances of well-known pieces without sound, participants were able to perform the task in short-term retention. It is also interesting to note that the benefit of using sound for learning here was restricted to concurrent sonification; provision of a sonified demo alone did not improve performance relative to control.

Chapter 6

Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning and listening for extended retention

A version of this chapter is peer-reviewed and published under the same title: Dyer, J., Stapleton, P, and Rodger, M. *Experimental Brain Research* (2017), doi: 10.1007/s00221-017-5047-8

6.1 Abstract

An important question for skill acquisition is whether and how augmented feedback can be designed to improve the learning of complex skills. Auditory information coupled to learners' actions, movement sonification, can enhance learning of a complex bimanual coordination skill, specifically polyrhythmic bimanual shape-tracing. However, it is not clear whether the coordination of polyrhythmic sequenced movements is enhanced by auditory-specified timing information alone or whether more complex sound mappings, such as melodic sonification, are necessary. Furthermore, while short-term retention of bimanual coordination performance has been shown with movement sonification training, longer term retention has yet to be demonstrated. In the present experiment, participants learned to trace a diamond shape with one hand while simultaneously tracing a triangle with the other to produce a sequenced 4:3 polyrhythmic timing pattern. Two groups of participants received real-time auditory feedback during training: Melodic Sonification (individual movements triggered a separate note of a melody) and Rhythmic Sonification (each movement triggered a generic percussive sound); while a third Control group received no augmented feedback.

Task acquisition and performance in immediate retention were superior in the Melodic Sonification group as compared to the Rhythmic Sonification and Control group. In a 24-hr retention phase, a decline in performance in the Melodic Sonification group was reversed by brief playback of the target pattern melody. These results show that melodic sonification of movement can provide advantages over augmented feedback which only provides timing information by better structuring the sequencing of timed actions, and also allow recovery of complex target patterns of movement after training. These findings have important implications for understanding the role of augmented perceptual information in skill learning, as well as its application to real-world training or rehabilitation scenarios.

6.2 Introduction

6.2.1 Information pickup in motor skill learning

Skilful control of human movement and the performance of learned motor skills are based in large part on the pickup and use of task-relevant perceptual information from the world via the senses. Generally speaking, availability of more precise, or higher-acuity perceptual information can enable finer control of movement, and more effective task performance (Mechsner et al., 2001; Todorov, Shadmehr, & Bizzi, 1997; Wilson et al., 2005). This improvement in performance can be afforded by the provision of concurrently-presented augmented feedback, which tracks some property of movement, (for example, deviation from a desired trajectory, see: Sigrist, Rauter, Riener and Wolf, 2013b), and feeds it back to the mover via sound in real time. The availability of higher-quality information allows enhanced performance-monitoring, by making errors more salient and correctable, and good performance more recognisable. When deployed in a motor skill learning scenario, i.e. a situation which entails deliberate repeated practice with the aim of improving performance on a given task (Ericsson et al., 1993), augmented feedback can lead to better learning outcomes than would be possible in the absence of feedback (Sigrist et al., 2013a). Traditionally, research into the effectiveness of augmented feedback has employed the visual modality, relying on continuously-updated visual displays to provide information for task performance (Kovacs et al., 2009; Vander Linden et al., 1993). 'Sonification', as it is used here, is live sound, controlled by learner movement[27], designed with the same aim of movement performance enhancement (Dyer et al., 2015; Effenberg, 2005).

A learner of a novel skill has access to abundant sensory information through his/her interactions with the environment. The role of the learner in this system is to refine and attune his/her attention to that information which is most relevant for effective completion of the task at hand (E. J. Gibson, 1969, 1988). Augmented perceptual information (as provided

[27] Sonification is more extensively studied in the broader context of auditory information display. For examples, see Dubus and Bresin (2013); Hermann, Hunt and Neuhoff (2011).

by concurrent feedback) can speed up this ‘education of attention’ by highlighting specifically the information which is relevant for the task, i.e. information for the control of the body or a tool. The learner’s action produces *more* useful information when he/she is immersed in an augmented feedback system. The ‘loop’ encompassing perception and action is thereby strengthened, and measurable improvements in performance come more quickly (Chang, Chang, Chang Chien, Chung, & Hsu, 2007; Kovacs & Shea, 2011; Todorov et al., 1997).

Chapter 5 reported an experiment showing a lack of a guidance effect in motor skill learning after practice with movement sonification. However, it is not simply the case that sonification as augmented feedback per se produces better, longer-lasting learning; rather there are structural features of some learning environments which make a guidance effect a less likely issue. In most cases, sonification shares these features. To elaborate, it has been noted that visual augmented feedback, when it is displayed as a live graph, represents a transformation of the information required to perform the task in the absence of feedback (Kovacs et al., 2010; Wilson, Snapp-Childs, Coats, et al., 2010), meaning that the skill which is learned is how to perceive and control the display, rather than the kinematics of limb movement. In contrast, feedback information which is untransformed, or that which does not alter the dynamics of the perception-action task at hand, produces performance increments which are resistant to the removal of feedback (Chiou & Chang, 2016).

Skilful performance is a product of attending to the right information for the immediate task (Wilson, Snapp-Childs, Coats, et al., 2010). It follows then, that effective augmented feedback should not transform the information which is necessary to perceive in a naturalistic performance scenario (i.e. without feedback). In fact the primary role for feedback (if the goal is learning which is not dependent *on* feedback) should be to highlight relevant features of task-intrinsic sensory information, i.e. to technically be redundant. Vinken et al. (2013) stress that sonification as augmented feedback should share a strong temporal-structural correlation with intrinsic information sources in order to take advantage of multisensory integration processes.

The experiment reported in Chapter 5 showed that non-abstracted, redundant sonification of movement enhances performance in a bimanual motor skill (4:3 rhythmic shape-tracing – the same task used in the current experiment). This task involved tracing two regular shapes on a workspace with the index finger of both hands concurrently (a triangle for the left hand and a diamond for the right). Learners were required to make regular movements between shape corners so as to produce an inter-corner bimanual timing ratio of 4:3. In one experimental condition, fingertip arrivals at corner zones were sonified – each producing an enveloped burst of pure tone, mapped to a specific pitch. When performed correctly, one full sequence produced a simple melody. Prior to each practice trial, learners were shown a visual demo animation of correct performance, sonified with the same mapping. Practice trials thereby became an attempted musical performance, as learners tried to match both the visual and auditory elements of the demo, and "play" the task correctly. As expected, when live sonified feedback was withdrawn, learners maintained the high level of performance they had reached by the end of practice (however this effect was only short-term, disappearing after 24 hours).

Although this result fits comfortably within the context of other recent work on sonification and the guidance effect (Chiou & Chang, 2016; Danna, Fontaine, et al., 2015; Heitger et al., 2012; Mononen et al., 2003), it seems likely that the explanation for these findings can be found in the structure of the augmented information provided, rather than the sensory modality (audition vs. vision). In the previous experiment, the onset and duration of musical tones was roughly matched to the movements of the learner, meaning that sonified feedback information was coupled to movement performance, without any transformation or abstraction from the movements themselves. When sonification was available, control of movement was enhanced. Sonification effectively directed learner attention towards a single set of perceptual events (fingertip corner arrivals), which were perceptually available in the same form as an intrinsic part of the motor task.. Learner performance was resistant to the removal of feedback because the information provided by feedback was perceptually congruent with task-intrinsic information, which was itself attended to during practice – a requirement for learning which generalises beyond feedback conditions (Sigrist et al., 2013a).

6.2.2 Potential benefits of a richer melodic mapping

Temporal control of human action is much improved when the action in question produces sound, or can be coupled with an external rhythmic stimulus (Kennel et al., 2015; Repp & Penel, 2002). This comes as no surprise, given that the temporal resolution of the auditory system is much finer than that of the somatosensory system (Hirsh & Watson, 1996; Tinazzi et al., 2002) and finer perceptual acuity affords more acuity in the control of associated action (Pizzera & Hohmann, 2015). Skilful performance in the 4:3 shape-tracing task is at least to some extent predicated on fine temporal control of action. It may then be the case that any kind of auditory information which shares the same structural features as those used in that experiment could produce similar performance benefits. In other words, making the task a melodic, conventionally musical task (as the previous experiment did), may not be essential or even necessary for movement performance enhancement. The same benefits might emerge if shape corner arrivals were sonified with identical bursts of sound, devoid of any corner-specific variation in pitch. This is one of the questions which the current experiment is designed to answer, by comparing a melodic sonification condition with another sonification condition in which corner arrivals are sonified using short, percussive bursts of white noise. If this “Rhythmic Sonification” condition produces performance benefits relative to control, then it might suggest that action-coupled sound per se can be sufficient for performance enhancement in rhythmic, continuous tasks such as this. Alternatively, there could be a performance advantage only in the “Melodic Sonification” condition, which would indicate that there is some value to giving such tasks a melody. Either of these findings may prove useful in the ongoing search for valid, experimentally-based justifications for the design of sonification action-sound mappings, a lack of which has been identified in the current literature (Dyer et al., 2015).

There is a way of conceptualising the task-sonification mapping for this task which suggests a benefit for melody over rhythm, based on information specificity. In this bimanual task, each corner of each shape is assigned a different tone. Played in order, they produce the melody heard in the demo presentation. This means that mistakes in the ordering of bimanual movements effectively stand out as incorrect. A move out-of-order means a note out-of-order in the increasingly familiar melody, and this mistake can be corrected on the next cycle. In the proposed Rhythmic Sonification condition, the salience of ordering errors is unlikely to be as great. If learners utilise this melodic strategy to learn the correct ordering of hand movements, it should manifest as a faster rate of learning in the proposed Melodic Sonification condition than the Rhythmic Sonification condition. If so, this result could serve as evidence-based justification for the use of melody in sequential motor skill learning with sonification. If there is a need for task-related actions to be performed in a specific order (examples of such tasks might include starting a car or performing a deadlift), a sonified training regime could be devised which incorporates the ordering of actions into a melody.

6.2.3 Prolonging motor skill retention through musical listening

A central concern in motor skill learning is the retention of good performance beyond the time period immediately following the practice phase. Delayed retention tests are essential to determine whether improved performance is a short-term effect, or can reasonably be called *learning* (Salmoni et al., 1984; Soderstrom & Bjork, 2015). Retention tests beyond the day of practice are somewhat rare in the published literature on sonification for motor skill learning. Konttinen, Mononen, Viitasalo and Mets (2004) provide one example of long term retention in the skill of rifle shooting. Gun barrel movements (relative to a target bullseye) were sonified with a sine tone, with the intention of training rifle stability and, indirectly, better accuracy. Higher scores were observed in learners who practiced with sonification relative to control in retention tests delayed 2, 10 and 40 days from practice. Conversely, learners in the previous experiment reported here showed a distinct decline in performance 24 hours following practice. The cohort who practiced with sonification of movement were statistically indistinguishable from control participants only a day later. Many participants in the sonification condition reported being unable to remember the melody from the previous day's practice, and claimed that this was the reason for their perceived poorer performance (no feedback was given). One solution which could remedy the 24-hour drop-off in performance would be to allow learners to listen to a replay of the sound of perfect performance before a delayed retention test – a strategy which has been effective in piano training (Lahav, Katz, Chess, & Saltzman, 2013).

Music perception is far from a passive process in which sounds are processed in isolation, with no relation to action. In fact, the fMRI activation crossover between perception of music and perception/performance of movement is substantial, especially for learned music (Bangert et al., 2006; J. L. Chen, Rae, & Watkins, 2012; Lahav et al., 2007; Lotze et al., 2003). This effect likely applies beyond the domain of learned music to action-relevant sound more generally (Cesari et al., 2014). Stienstra, Overbeeke and Wensveen (2011) sonified foot pressure on different regions of a speed-skater's boot with dynamically filtered and pitch-modulated pink noise. The description of how the skater learned to perceive movement from a soundscape through interactive practice (while those who had never used the system could not) effectively illustrates how there can be abundant action-relevant information in sound for a listener who is skilled enough to perceive it (for a more detailed description of this experiment and the findings, see section 3.4. For the skater in this example, listening to the sound of the system became the direct perception of an ongoing action, not dissimilar in terms of action-specification from a first person point-of-view visual demonstration. The skilled listener, i.e. one who has learned the mapping between movement and sound in an interactive sonification environment, should be able to perceive a great deal of useful information from a replay of the sound of perfect motor performance. For the current experiment, it is hypothesised that on a 24-hour retention test, participants will show improved motor performance following a short listening period (pre-recorded sound only, no visual presentation or live movement sonification). This finding would show that there is potential for sonification-enhanced training to be 'refreshed' by listening to the sound of good performance. This could be useful in sports training where a skill trained to a high level of performance in a sonification lab may start to decline over time. Rather than expend time and effort returning to the lab, the athlete could instead listen to a recording of their sonified performance when they were at their best. This alone could be sufficient to allow good performance to re-emerge, as the auditory demonstration would effectively be (some part of) the holistic perceptual experience of good performance, which can then be replicated (Rosenblum et al., 2016). Because of the ubiquity of personal headphones, this strategy could even be enacted in scenarios where a visual replay might be beneficial, but impractical, e.g. on a golf course or on the field (Agostini, Righi, Galmonte, & Bruno, 2004; Murgia et al., 2012).

The aim of the present experiment is to investigate two separate issues of sonification for motor skill learning. The first is whether there is a specific benefit to using a musical, melodic sonification in ordered, sequential tasks such as the task to be learned here - as opposed to a sonification strategy which provides non-melodic, rhythmic information only. To address this, three experimental conditions are proposed. A "Melodic Sonification" condition, in which movement events are sonified using a selection of notes (which together, and in order, form a melody), and a "Rhythmic Sonification" condition, in which the same movement events are sonified using identical bursts of white noise. These are compared to each other and a control condition, in which no auditory feedback of any kind is provided. To be sure of eliminating audition from the perception-action task, participants in this condition hear constant, task-irrelevant pink noise at a comfortable volume. The second issue is the extension of good performance beyond an initial retention test on the day of practice. Participants in all three conditions are tested twice on day two. The second such test is preceded by an auditory playback of the demonstration (Melodic or Rhythmic depending on the condition). The control condition will complete two identical retention tests on day two, with no demo presentation of any kind. This is intended to address the issue of a potential practice effect due to multiple retention tests in a short time period. By comparing the change in performance on the control condition with the other two conditions (post replay), it should be possible to reasonably assert whether the change was due to repeated performance i.e. practice, or from the sonic replay.

6.3 Methods

Participants

An opportunity sample of 60 participants (39 female, 21 male) was recruited from a pool of undergraduate Psychology students, postgraduate researchers and staff in the university at which the experiment was conducted. Undergraduate students received course credit for their participation where applicable. Only right-handed participants were recruited, as confirmed by administration of the Edinburgh Handedness Inventory (Oldfield, 1971). Handedness scores did not differ significantly between the three experimental groups, as confirmed by ANOVA: $F(2, 59) = .260$, $p = .772$).

Participants were asked to report any musical experience or participation in dance activity, no matter how small or long ago. No professional musicians or drummers were included in the sample. In the control condition there were eight participants with some musical experience (mean 8 years, S.D. = 3.91), four of whom were currently in some way involved in music (e.g. recreational players/learners) and one participant who was a regular dancer. In the Rhythmic Sonification condition there were eight participants with musical experience (mean 9 years, S.D. = 3.16), three of whom were currently involved in music, with one dancer. In the Melodic Sonification condition there were 9 participants with musical experience (mean 6.33 years, S.D. = 3.28), two of whom were currently involved in music.

Informed consent was obtained from all individual participants included in the study. Ethical approval for this study was granted by the School of Psychology Ethics Board at Queen’s University, Belfast.

Materials and Apparatus

Materials and apparatus used in the current experiment are identical to those reported in Chapter 5. In this section, only features of the experiment which differed are reported.

Feedback

Three kinds of feedback were provided in the current experiment. Concurrent Melodic Sonification and Rhythmic Sonification were available to participants during practice trials in those experimental conditions. Terminal feedback in the form of a graph of performance was provided to all participants (including those in the control condition) following every trial. All three kinds of feedback are based on the same movement events: index fingertip arrivals at corner zones of the shapes.

When the index finger of either hand entered a zone defined around a shape corner[28], a trigger was produced in *Max/MSP*. In the Melodic Sonification condition, this produced a note from the melody shown in Figure 6.1.

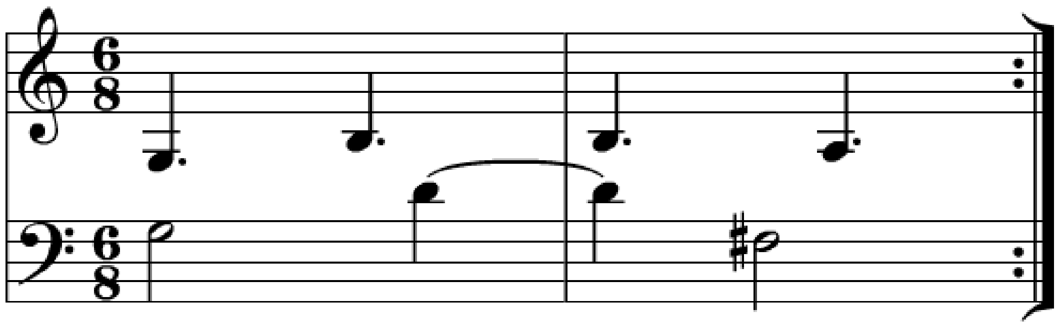

Figure 6.1: The demonstration melody presented in the Melodic Sonification condition and produced by correct performance by participants in the same condition. When movements were sonified, right hand movements produced notes from the upper row, and left from the lower.

Notes were synthesised using a version of the Karplus-Strong string synthesis procedure[29], which is based on a physical model of a plucked string (Karplus & Strong, 1983). When played correctly, each note activates with an initial high intensity and decays

[28] Corner zones were defined in two axes (*x*&*y*), thus forming a square of 3x3 cm centred on each corner.

[29] Patch available online at: bit.ly/2hnBGFv

roughly exponentially to total silence in approximately 1000ms. An additional velocity mapping was included in this sonification procedure, whereby movements between corner zones which were excessively slow would prolong the length of the initial high-intensity impulse of the upcoming note while also reducing the 'brightness' of the sound. This feature was triggered when the maximum recorded velocity between two corners fell below a threshold of .17 m/s, after which the decaying current note could potentially be just audible beyond the onset of the next. In perception-action terms, this meant that when a participant did not stop at corner zones but rather continued in a slow, continuous fashion, consecutive notes would appear to blend together. Discrete movements (with peak velocity > .17m/s) between corners produced a short-duration initial impulse with a 'bright' quality, which meant that consecutive notes were perceptually discrete. No specific instructions were given to participants regarding how to move between corners in any of the three conditions. A link to a video recording of this sonification procedure can be found in Appendix A.

The demo animation for the Melodic Sonification condition was sonified as if movements were performed with an acceptably high velocity, i.e. with no extra duration. During both the demo presentation and live movement sonification, tones produced by movement of the left hand were panned to the left channel of the headphones and vice versa.

In the Rhythmic Sonification condition, corner arrivals were sonified with bursts of white noise. Loudness of the burst was modulated by an envelope function which reduced loudness in a roughly exponential fashion until silence at 350 ms after onset. The demo animation was sonified with the same sounds. As in the Melodic Sonification condition, left hand-produced sound was panned to the left and vice versa.

Every trial in the practice stage was concluded with the presentation of a line graph of performance (see Figure 3), showing raw ratio data for the previous trial relative to perfect performance. Throughout each trial, inter-corner intervals for the right hand were calculated and compared to the previous inter-corner interval on the left hand to produce a ratio. The ideal right-to-left ratio (3:4) was displayed on the graph as a green horizontal line across the centre of the screen. Participant-produced ratios were displayed as dots connected by a line.

Procedure

Participants were each randomly allocated to one of the three experimental conditions (N = 20 in each). The experiment proceeded in seven stages for each participant.

Stage 1: Familiarisation

As in the previous experiment, the visual demo animation was played twice without sound prior to the practice phase. This animation showed the corners corresponding to the apparatus shapes lighting up in sequence, demonstrating the spatiotemporal requirements of the movement task. The top corners of both shapes lit up simultaneously. Corners on the left (triangle) then lit up once every 1000 ms, while corners on the right (diamond) lit up every 750 ms, both in an anticlockwise direction. A full cycle thus lasted three seconds exactly. Every play of the demo (including in the later practice phase) consisted of three cycles (totalling 9 seconds in length). For familiarisation, it was played twice (18 seconds total). Participants were then given time to attempt to produce the movements they had observed in the demo without their performance being recorded (approximately 15 seconds). No sound was presented during familiarisation.

Stage 2: Pre-test

Participants in all conditions completed a single trial under control conditions to ensure equality of performance, on average, at the outset. This trial was performed while listening to pink noise through headphones during the demo and movement phases. The demo was played once (3 cycles, 9 seconds total), then participants were given 26 seconds in which to produce the movements of the task. During this time, they were required to produce continuous cycles on the shapes. No feedback (auditory or graphical) was provided on this trial.

Stage 3: Practice

The procedure for practice trials did not differ between groups except in terms of the auditory information available. Practice trials commenced with a play of the demo, followed immediately by a 26-second recorded movement phase and concluded with the presentation of the terminal feedback graph. Fourteen practice trials were completed in total.

In the Melodic and Rhythmic Sonification conditions, the demo and participant movements were sonified as described earlier, such that perfect performance by participants produced exactly the same sound heard during the demo. In the control condition, constant task-irrelevant pink noise was heard throughout the demo and movement phases.

Stage 4: Short-term retention (post-test)

Following a break of five minutes, a 26-second retention test was administered without a demo presentation or any form of feedback (neither sonification nor graph). During this time, participants were invited to perform the task to the best of their ability, while listening to constant pink noise.

Stage 5: 24-hour retention

The following day, participants returned to the lab to repeat the retention test from the day before exactly.

Stage 6: 24-hour post-replay retention

Participants performed another retention test under the same conditions as previously described. However, prior to movement, participants in the two sonification conditions heard the sound produced by perfect performance of the task according to their condition (i.e. those in the Melodic Sonification condition heard the melodic demo, and so on). The sound of the demo was played twice, without any accompanying visuals (6 cycles, 18 seconds of sound total). Participants in the control condition did not hear nor see any task-related information prior to their test.

Stage 7: Transfer

A transfer test was conducted here to ascertain whether there might be differential transfer of learning between feedback conditions. The test involved the switching of the shapes. The triangle was placed on the right, and the diamond on the left. The task goals were the same (4:3 bimanual rhythmic shape tracing), only the apparatus was mirrored.

Analyses

The main measure of performance in the current experiment is bimanual timing ratio error, which was calculated as described for the previous experiment. Analysis in this experiment focuses mainly on detecting potential benefits of sonification (either type) relative to the Control condition. ANOVA are employed to test for differences in performance between feedback conditions at relevant time points in the experiment. Rates of learning for each condition were examined by fitting curves to performance data from individual participants in the practice stage (trials 1-14) and comparing mean slope values. Some of the retention data were subjected to a confidence interval-based statistical test of non-inferiority, as described in the previous chapter.

6.5 Results

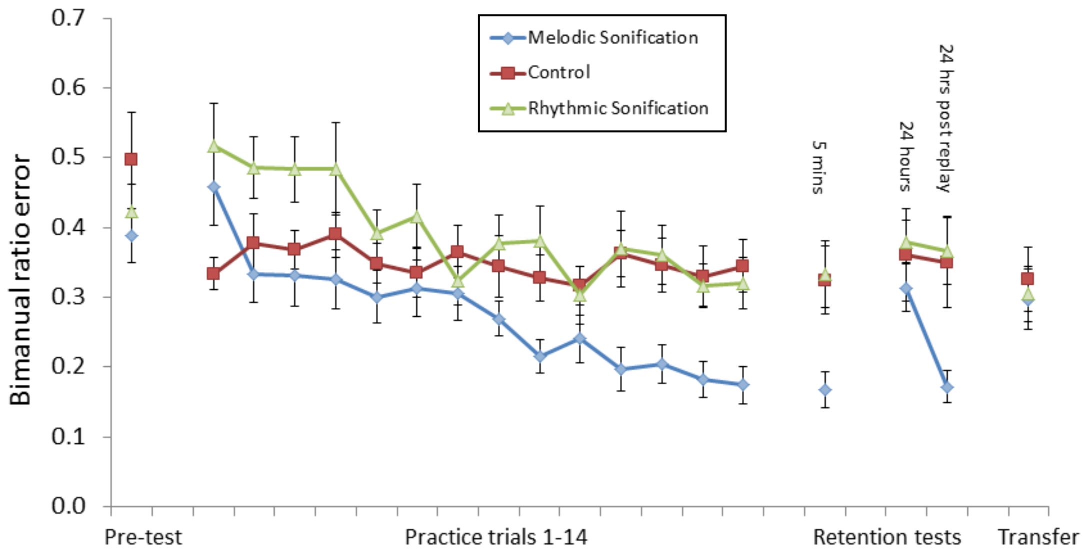

Figure 6.2: Rates of average absolute bimanual ratio error per condition in pre-test, practice, retention testing and transfer testing for each of the three training condition groups. A score of 0 represents perfect performance. Error bars are standard error.

6.5.1 Pre-test

A one-way ANOVA on data from the pre-test revealed no significant effect of feedback condition on scores: $F(2, 58) = 1.090$, $p = .344$.

6.5.2 Practice trials 1-14

A mixed ANOVA across all acquisition data (trials 1-14, all participants) with feedback condition as a between-groups factor and practice trial as a repeated measures factor revealed a significant main effect of trial: $F(6.059, 260.525) = 7.867$, $p < .001$, $\eta^2 = .062$ but not of group: $F(2, 43) = 2.589$, $p = .087$. A significant interaction effect of trial*group was found: $F(12.117, 260.525) = 2.818$, $p = .001$, $\eta^2 = .044$. Pairwise comparisons of group performance were performed on data from the final practice trial (14) to test for the effect of sonification on performance by the end of the practice stage, while sonification was still available (alpha was set at .016 – Bonferroni correction for three comparisons). These indicated that participants in the Melodic Sonification condition performed significantly better than those in the Rhythmic Sonification condition: $p = .004$, Cohen's $d = 1.06$; and the Control condition: $p = .001$, Cohen's $d = .959$. Performance did not differ significantly between the Control and Rhythmic Sonification conditions: $p = .593$.

Rate of learning is a central concern here, therefore curves were fitted to performance data from practice trials 1 to 14 as described in section 4.4.4. Holm-Bonferroni-corrected *p* values are reported for this analysis. Slope values obtained were not significantly greater than 0 in the Rhythmic Sonification condition: $t(19) = 2.026$, $p = .029$; nor in the Melodic Sonification condition: $t(19) = 1.626$, $p= .180$; nor the Control condition: $t(19) = 1.546$, $p = .180$. Between group comparisons revealed no significant effect of condition mean slope values: $F(2, 59) = .123$, $p = .885$.

6.5.3 Retention testing

The mean of the difference scores for the Melodic Sonification condition between trial 14 and the 5-minute retention test was .01, with a 90% CI of [-.01, .03]. The upper confidence interval is .03, which falls below the non-inferiority interval of .085. A *p* value is provided for this test by performing a one-sided, one sample *t* test on difference scores relative to the non-inferiority interval, .085: $t(19) = -6.98$, $p < .001$. It can therefore be inferred that the improved performance relative to control observed in the Melodic

Sonification condition did not diminish without the presence of sound (see Figure 6.2). At the first 24-hour retention test, ANOVA revealed no significant effect of feedback condition on error scores: $F(2, 59) = .588$, $p = .558$. The improved performance in the Melodic Sonification group was not evident after 24 hours.

Following the sonic replay, a significant effect of feedback condition on error scores was again detected: $F(2, 59) = 5.208$, $p = .008$. Post-hoc *t* tests showed that this effect was driven by significantly lower error scores ($\alpha = .016$) in the Melodic Sonification condition relative to Control ($p = .010$) and Rhythmic Sonification ($p = .006$). Performance did not differ significantly between Rhythmic Sonification and Control ($p = .830$). A further test of non-inferiority relative to performance at trial 14 was performed on data from the Melodic Sonification condition following the replay. The mean of the difference scores between these two points was .002, with a 90 CI of [-.03, .03]. The upper CI falls within the .085 threshold for non-inferiority ($t(19) = -5.319$, $p < .001$), therefore it can be inferred that performance in the Melodic Sonification condition was statistically no worse after 24 hours and a sonic replay than at the final practice trial.

To test for a potential practice effect of repeated performance on retention test scores, data from 24-hour retention (pre and post replay) from the control condition were subjected to a paired samples *t* test. No significant change in performance between these two performance tests was evident. The mean score at the first of the two tests was .362 (SD = .291), at the second .350 (SD = .284); $t(19) = .553$, $p = .587$.

6.5.4 Transfer testing

ANOVA revealed no significant effect of feedback condition on performance in the transfer test: $F(2,59) = .128$, $p = .880$.

6.6 Discussion

6.6.1 Performance benefits with the use of melody

In this experiment, sonification of movement which employed a melodic mapping was more effective for motor skill learning than a similarly-structured sonification which consisted only of basic temporal (rhythmic) information. By the end of practice, average bimanual ratio error was significantly lower in the Melodic Sonification condition than both the Rhythmic condition and the Control condition. The Rhythmic Sonification, which was expected to improve performance by more clearly specifying the timing of required and produced participant movements, did not improve performance relative to Control. This result indicates that the improved performance observed in the sonification condition in the previous experiment was not only due to action-sound coupling, but an action-sound coupling which produced a meaningful melodic pattern.

That Rhythmic Sonification of movement did not produce an improvement in performance at all relative to Control is somewhat surprising. From the sensorimotor timing and motor control literature, it is clear that actions performed in the presence of a sonic metronome, or intrinsically sounding actions can be performed with greater temporal accuracy than similar actions performed in silence (Kennel et al., 2015; Repp & Penel, 2002). Good performance on this task (which shares some fundamental characteristics with classic polyrhythmic coordination tasks, see Summers et al., 1993) is at least partly dependent on the fine temporal control which is afforded by action-sound coupling. The timing structure of the auditory information provided in this Rhythmic Sonification condition was essentially the same as that provided in the Melodic Sonification condition and the sonification condition in the previous experiment. Sound events were coupled directly to movement events and the resulting auditory information was not transformed or abstracted from the basic underlying kinematics of the task - which should have made it directly useful for the coordination of action (Chiou & Chang, 2016). However, it is important to note that this task is more complex than the tasks used in the bimanual tapping paradigm (D. M. Kennedy et al., 2013; Klapp, Nelson, & Jagacinski, 1998). Most of the research which informs the above assumption (that attaching any sound to movements will improve rhythmic bimanual coordination performance) comes from uni- or bimanual finger tapping on static force plates, positioned directly beneath the hands. No such research into the effect of sonification on polyrhythmic bimanual coordination in more complex tasks yet exists (to the author's knowledge). The requirements of this task include perceptually conflicting features such as movements in different directions, amplitudes, using non-homologous muscles and at different times - all of which are known to increase coordination difficulty (Shea et al., 2016; Swinnen & Gooijers, 2015).

In this task, the benefit of melody can be concretely conceptualised in terms of the richer, more useful information it provides in addition to inter-movement interval durations. Attaching specific notes to individual corner zones clearly specifies the order in which movements must be performed. Mistakes in the ordering of movements are reflected in very salient mistakes in the unfolding melody produced during the cycle. On the following cycle, this can be corrected. Ordering mistakes are reflected in the main measure of bimanual ratio error, as movements out of order are movements which come too soon or too late, and thereby affect the bimanual timing ratio. Conversely, with the Rhythmic Sonification, it is possible to produce a rhythm very close to that presented in the demo, while still making ordering mistakes in the execution of the task (aside from the lateral stereo panning, which would reveal the error, however this is likely much perceptually salient than mistakes in a melody). It was predicted that the informational benefit for melody in this task would manifest as a faster rate of learning in the Melodic Sonification condition, however, analysis of slopes obtained from fitting curves to the data of individual participants did not proffer any evidence of this. It is possible that participant performance displayed too much trial-to-trial score variability in (primarily during the early stages of practice) to enable the fitting of truly representative curves, which could accurately describe rate of learning. This may have hindered the efficacy of the analysis technique, leading to the null findings reported here.

Despite better performance in the Melodic Sonification condition by the end of practice, it is possible that, given more time, participants in the Rhythmic Sonification condition could have learned to use the spatial, stereo-panned information to better control the ordering of their movements, given that ordering is *technically* specified in the audio information available. Their performance may eventually have equalled that seen in the Melodic Sonification condition, however with the necessary specifying information being so subtle, it is unclear how long such improvement would take (Carello, Wagman, & Turvey, 2005).

6.6.2 Learning and refresh with musical learning

In the five-minute retention test, with all feedback removed, task performance was unchanged. The most interesting result from the first retention test is that the enhanced performance seen in the Melodic Sonification condition was not dependent on the presence of live feedback. Participants were no worse at the task without sonification. This is a replication of the main finding from the previous experiment, showing the absence of an early-retention guidance effect. Similar results have been reported by Ronsse et al. (2011) and van Vugt and Tillmann (2015) who directly sonified sequential actions and observed maintenance of good performance after the removal of sound.

On the 24-hour retention test, the improved performance in the Melodic Sonification condition had disappeared; no effect of condition on performance was detected at this point. This was expected, and lines up with an identical previous finding. In the previous study, participants frequently reported that they were unable to remember the melody, and blamed that for the decline in their performance. It was these reports and the notion of perception and action as a holistic process which inspired the attempt in the current study to prolong retention with the use of a sonic replay. The behavioural and neural crossover between perception of sound and action production in musical skill is very well-established (Lahav et al., 2007; Lotze et al., 2003; Taylor & Witt, 2015). In piano learning, retention of a learned sequence of notes can be enhanced by motionless listening to the sound of correct performance (Lahav et al., 2013). A major strength of sonification as a vehicle for the delivery of augmented feedback information is that it can very easily transform non-musical tasks (with more abstract goals than the production of music) into musical tasks with features akin to traditional musicianship.

Following only 18 seconds of listening (two plays of the demo), differences between feedback conditions became evident again. Participants in the Melodic Sonification condition performed the task with lower rates of error than both Control and the Rhythmic Sonification condition. In fact, a statistical test of non-inferiority showed that performance in the Melodic Sonification condition at this point was not any worse than it had been at the very end of the practice stage on the day before. Note that no augmented feedback whatsoever was available to participants on this post-replay test. This indicates that, despite the poorer performance shown on the initial 24-hour retention test, the motor skill did in fact remain in the repertoire. In this experiment, melody represented the key to unlocking good motor performance after it had waned.

Beyond this experiment, the broader application of this finding is that novel motor skills can be trained with sonification and that good performance can be refreshed by listening to its ideal sound. However, no corresponding benefit of listening to a sonic replay was evident in the Rhythmic Sonification condition. Again, this is likely due to the information provided by sound, and the degree to which it specifies the movements of the task. The sound of the melody specifies the ordering of the movements of the task, but crucially, only for those participants who are skilled enough to perceive the action-relevance of that information (Carello et al., 2005; Steenson & Rodger, 2015). As in the speed-skating example of Stienstra et al. (2011), there is rich, action-relevant information in sound which can be used to directly perceive the interactive task as a whole (both perception and action components) when listened to by a skilled individual. A demonstration of task performance through feedback which is more abstract, or does not as precisely specify the interaction would likely not be as effective for refreshing motor performance. For example, it is difficult to imagine that viewing a Lissajous figure in motion after good performance in a bimanual coordination task has extinguished would result in the re-emergence of good performance in no-feedback retention. The information provided by such displays is abstracted, transformed from the basic kinematics, but most importantly, generic. The Lissajous figure does not precisely enough specify individual limb kinematics to allow limb movement to be perceived from it. In a similar vein, the sound of the Rhythmic Sonification demo in the current experiment does not clearly enough specify the ordering of the movements sonified for the sound alone to be useful in the same way as the melody - the sounds used are too generic and are not easily distinguishable as caused by specific limbs in specific motion.

This finding has been implicitly acknowledged in other work which has used sonification. For example, Ronsse et al. (2011) opted for two-tone sonification (successful performance of their task produced four evenly-spaced tones, in what sounded like a galloping rhythm) rather than one, which would have been as generic as the Rhythmic Sonification in the current experiment. Four evenly-spaced tones which are all the same would have been much more difficult to use to coordinate a difficult *bimanual* skill. It might be that the task is complex enough to warrant two-tone sonification. Melodic information as simple as variation in pitch between two beeps (for two hand orientations) may be sufficient to more clearly specify task requirements, as it is in Ronsse et al., however more complex motor skill sonification may require a correspondingly elaborate melody, or otherwise more complex auditory information.

Performance in the transfer test did not differ between experimental conditions. Despite the fact that participants in the Melodic Sonification condition were able to improve their performance on the main task to a level equivalent to the previous day, this did not affect performance on a mirrored version of the task.

6.6.3 Motivation

The effects of intrinsic motivation and task engagement are often not explicitly considered in tests of sonification for motor skill learning, despite the well-known mood enhancing (and consequent task performance enhancement) effect of enjoyed music (Schellenberg & Hallam, 2005). Melodic musical phrasing has been characterised as a unified perceptual whole, analogous to how discrete patches in the visual field are perceived as whole objects despite being partially obscured by another object in the foreground (Bregman, 2001; Bregman & Campbell, 1971). Successful production of a melody on a musical instrument can be intrinsically rewarding, as much as (or perhaps more than) resolving an ambiguous object in the visual field (Blood & Zatorre, 2001). Sonifications for motor skill learning, on the other hand, very often employ aesthetically impoverished sound (such as continuous sine tones, or pink noise) directly mapped to numerical values - values which it is assumed that listeners will be able to pick up via sound and understand in the context of their motor performance (Dubus & Bresin, 2013). Sonification systems which use these kinds of mappings are not intrinsically rewarding or purposeful to act within; they are intended only to inform intellectually, while the quality of experience is neglected (see section 3.1.7). However if a listener is not sufficiently motivated to be attentive, there could be negative implications for the pickup of information, and therefore, task performance. In broad terms, it is important to remember that perception is an act which an agent must be motivated to perform.

6.6.3 Caveats and conclusions

It seems clear from the results presented here that if a sonification system were to be devised for training a sequential, ordered task, then the use of melody would be well justified. Caution must be taken when applying this design strategy, however, as it is likely that not all melodies are equally useful. "Melody" is an extremely broad and subjectively-defined term, encompassing potentially infinite possibilities for the combination of different notes. The melody composed for the current experiment was deliberately designed to sound pleasant; it is played in a major key, easy to remember and sounds 'complete' at the end of each cycle. It could be speculated that what might technically be described as a melody, but designed with all the opposite features (minor key, discordant and forgettable) would not be as useful in a sonification for motor skill learning.

The current study presents the benefits of melodic sonification for learning of a novel motor skill. The use of melody in the practice phase allowed participants in that condition to reach significantly lower error scores than control, and a sonification which used only rhythmic information. It has been argued that the main mechanism driving this effect is the extra information-structural variation available with the use of melody, and its ability to specify the ordering of the task, whereas purely rhythmic information does not allow for this. The secondary finding is that after performance has declined on a musical sonification-trained task, performance can be refreshed by listening to the sound of a perfectly-performed demonstration. Re-exposure to an augmented feedback-enhanced learning environment is not necessary if the skill is reconceptualised as a musical task. These findings have important implications for understanding the role of augmented perceptual information in skill learning, as well as its application to real-world training or rehabilitation scenarios.

Chapter 7

Sonic information design considerations: Information or knowledge as auditory feedback for motor skill learning

7.1 Abstract

Learning of new motor skills can be assisted by the provision of sonification as feedback - sound which is controlled by learner movement and fed back live. Many implementations of this technique define and sonify a variable of task performance, implicitly assuming isomorphism between measured improvement and motor skill learning. A more effective approach for learning may be to use sound to highlight task-intrinsic perceptual information. In the current experiment, a *metric of performance*, as measured by the experimenter, is fed back to participants as they learn a complex motor skill (4:3 bimanual coordination). In another condition, task-intrinsic events in the perceptual-motor workspace are sonified in a direct manner, using a melody. An additional group of participants practiced without sonification. Performance was superior in acquisition and short-term retention testing in the Melodic condition, however performance in the Metric condition did not decline as severely following a 24-hour delay. Following passive listening to the sound associated with perfect task performance, the Melodic condition improved to a level consistent with the previous day's superiority. A similar effect was not observed in the Metric condition. Results are discussed in the context of auditory feedback strategies in Psychological experiments and broader discussions around sonic information design.

7.2 Introduction

7.2.1 Sonification and motor skill learning

'Sonification' is the transformation of a data stream into sound so that its content might be experienced sonically. This is an approach championed by the 'Auditory Display' community, whose members seek to make datasets more accessible and computer systems more informative through the innovative use of sound (Hermann, Hunt, & Neuhoff, 2011). Sonification is also gaining acceptance in Psychology, where researchers have begun to test the utility of sound as performance feedback to be delivered to a learner of a motor skill (Boyer et al., 2016; Danna, Fontaine, et al., 2015). However fundamental questions about how sonic information should be designed to best enhance motor skill learning remain unanswered. The current chapter examines some of these questions in detail, first identifying a common style of sonification used in motor skill learning, then considering an alternative which could be more effective.

Motor skill learning is in part based on judgement of the success of a performed action in the context of task goals. This judgement is a product of successful pickup and use of task-relevant information via perception (Newell et al., 1991). In everyday learning scenarios, task-relevant information may be picked up through multiple sensory modalities simultaneously. Such information is said to be 'intrinsic' to the task, and available to any learner with the appropriate sensory apparatus. Additional, 'extrinsic', or sometimes, 'augmented' information can be delivered to the learner through some mediated system, for example a coach, graphical display, or speakers. Augmented feedback, when delivered concurrently with movement performance, can be an effective tool to enhance the acquisition of novel motor skills (Kovacs & Shea, 2011; Maslovat et al., 2009; Sigrist et al., 2013a).

Sonification in this context is augmented feedback delivered in the auditory modality - an additional sonic information channel through which movement performance can be listened to (Effenberg, 2005). In these systems, movement, or a change in how the task is being performed, is near-instantaneously reflected in the sound heard (usually with some imperceptibly small latency - a variable constraint of the computerised system(s) employed). Existing empirical research demonstrates the effectiveness of sonification as feedback in a range of tasks, including shooting (Konttinen et al., 2004), rowing (Schaffert & Mattes, 2015), pommel-horse (Baudry et al., 2006), handwriting (Danna, Fontaine, et al., 2015), coordinated finger-tapping (van Vugt & Tillmann, 2015), pointing (Schmitz & Bock, 2014) and bimanual coordination (Dyer, Stapleton, & Rodger, 2017; Ronsse et al., 2011). Sonification has also been successfully trialled in rehabilitation contexts, including for stroke (Scholz et al., 2015; Wallis et al., 2007), Parkinson's disease (Rodger, Young, & Craig, 2014) and graphomotor disorders (Danna et al., 2014).

7.2.2 Mapping movement to sound

While sonification has been practiced in the Auditory Display community for at least a few decades, its emergence as a modality of augmented feedback came more recently (Höner, 2011; Sigrist et al., 2013a). As such, there are still many open questions around the use of sonification for the enhancement of motor skill learning. The most central of these concerns the relationship between movement and sound - the mapping. Commensurate with the relative youth of the field, there is no unifying set of guidelines or accepted practices for how feedback for a task should be provided through sonification - which variables should be tracked and how they should be presented sonically (see Chapter 3). Sonification systems are generally bespoke, designed for the task of interest, and reports seldom include rationale for the sonification of a particular variable, or the use of a certain sound morphology. This leaves open the possibility of implementing sonification mappings which are sub-optimal, or even detrimental for learning.

For example, Baudry et al. (2006) used a simple form of sonification, an 'auditory alarm' to train performance on the pommel-horse. Their system produced an auditory signal (a pure tone) when the angle between the torso and legs deviated unacceptably far from 180° (straight form is desirable on the pommel-horse). This served as an indication that performance was deficient, and that a correction should be made. Other implementations have similarly indicated when motor performance was incorrect, and the extent of the error. Schmitz and Bock (2014) tracked positional error in a pointing task and mapped the resultant value to the pitch of a pure tone in a continuous fashion. Pointer movement towards the target altered the pitch of the tone, which was heard constantly, reaching a predefined target pitch when the participant was pointing directly at the target. Konttinen et al. (2004) used the same strategy in shooting, mapping the deviation between current barrel position and the centre of a bullseye to the pitch of a pure tone - such that the highest pitch was audible when the barrel was aligned with the centre of the target. In one of a set of experiments, Rosati, Oscari, Spagnol, Avanzini and Masiero (2012) sonified two dimensions of positional error in a visuo-motor tracking task. Error in the *x* axis was mapped to the amplitude and fundamental frequency of a synthesised vowel sound, while error in the *y* axis was mapped to the spectral content of the sound. This strategy was found not to be as successful as another which directly sonified the movement of the to-be-tracked target. In a lab-based rowing task, Sigrist, Rauter, Riener and Wolf (2013b) also presented two dimensions of positional error through sound. Vertical deviation from the desired oar trajectory was mapped to sound pitch, while horizontal deviation was mapped to inter-aural stereo panning. When compared to other conditions which had access to visual, haptic or no augmented feedback, the sonification condition was the only group not to show any effect of practice.

7.2.3 Problems with sonifying the measure variable

The examples related above, while drawn selectively from the literature, illustrate a prevalent strategy in the design of sonifications for motor skill learning. That is, the identification and display of a performance metric which has a certain known quantity (e.g. zero positional error, or perfectly straight form) if the task is performed correctly. The identity of this variable reflects how task performance is measured by the experimenter. In Sigrist et al. (2013b), the task is to perfectly reproduce a trapezoid-shaped oar trajectory.

Performance assessment (by the experimenter) requires continuous measurement of oar position deviations relative to the target in the *x* and *y* axes. Since these variables combined represent performance, it might seem that they are obvious candidates for sonification. If the scores they represent can be brought into an acceptable range, then the task is - *by definition* - being performed correctly.

7.2.4 Guidance effect implications

Two overlapping concerns arise upon consideration of this approach. First, feeding an experimenter's metric back to the participant means that the participant's primary task is to learn behavioural strategies to control that variable. While the development of such a strategy may result in 'correct' performance as measured by the experimenter, it may not necessarily be identical to motor skill learning – rather it is a strategy for control of the feedback system and may only be effective when sonification is present. The perception-action approach to motor skill learning explicated in Chapter 1 suggests that learning of a new skill is in large part characterised by 'education of attention' - movement towards the pickup of useful informational variables which support coordination with one's environment (Fowler & Turvey, 1978; Ingold, 2001; Jacobs & Michaels, 2007). This 'bottom-up' conceptualisation is in contrast to the classical view of motor skill learning as the construction of a 'motor program' or 'action plan' by way of internal processes (Chamberlin & Magill, 1992; Wolpert, Diedrichsen, & Flanagan, 2011). It is supported by a substantial body of literature showing a 'guidance effect' of augmented feedback (i.e. performance which is dependent on the presence of feedback, and immediate decline upon its withdrawal, e.g. Kovacs & Shea, 2011; Maslovat et al., 2009; Ronsse et al., 2011) and related research which demonstrates the potent effect of perceptual transformation on performance of complex motor tasks (Franz, Zelaznik, Swinnen, & Walter, 2001; Kovacs, Buchanan, & Shea, 2010; Mechsner, Kerzel, Knoblich, & Prinz, 2001; see also: Wilson, Snapp-Childs, & Bingham, 2010). In other words, motor skill acquisition requires learning how to pick up task-relevant information for immediate use, and is not purely a product of repeatedly moving correctly. The implication of this insight for sonification design concerns the guidance effect. 'Metric sonification' (for lack of better-established terminology), as described previously, implicitly frames task requirements from the detached perspective of the experimenter - in the form of measurement variables. As such, the available informational variables - with which the learner is encouraged to coordinate - may not correspond to the informational variables which would be crucial for performance in the absence of sonification. Learners may thus acquire a perception-action strategy which only works when feedback is present, leading to a guidance effect.

7.2.5 Burdensome intellectual elaboration

The second concern relates to the nature of the information made available by this style of sonification and how it differs from naturally-occurring, ‘ecological’ information. Ecological information, in the context of a motor task, is inherently temporal. Movement is required for its pickup, and it can specify the dynamics of the ongoing interaction - allowing direct perception of events in the 'perceptual-motor workspace' (Newell et al., 1991) and thereby, movement performance (Fowler & Turvey, 1978; Turvey et al., 1981). Effenberg (2005) suggests that sonification should work in the same way, i.e. “…if there’s no movement at all, there’s no sound.” (p. 53). In addition to requiring movement for pickup, ecological information does not have intrinsic content; its ‘meaning’ is predicated on the context of its continuous use in a task (van Dijk et al., 2015). This lack of 'content' is no impediment for a skilled perceiver who has learned how to make use of specific patterns of information through direct experience. In a piano-learning experiment, Lahav, Katz, Chess and Saltzman (2013) demonstrate that ostensibly 'passive' listening to the sound of an already-learned (short) piece can enhance its retention as a motor sequence. For such individuals, listening to piano music is not a task which requires intellectual work to extract 'content' - or a message - from an incoming auditory signal; what could this message *be* in any case? Ecological information, like the sounds produced by traditional musical instruments and other sounds whose production is constrained by the everyday physical laws enforced in the environment, can provide skilled listeners with the ability to perceive qualities of movement directly, e.g. the size and 'graspabilty' of objects from their sound spectra when dropped (Sedda et al., 2011), or the time-to-arrival of a vehicle from the same (Schiff & Oldak, 1990; see also: Steenson & Rodger, 2015; Carello, Wagman, & Turvey, 2005).

However, ‘metric sonification’ is often designed to have intrinsic content - or more specifically, to transmit a message about how the learner's performance measures up in terms of the experimenter's metrics[30]. In Schmitz and Bock (2014) for example, the sound heard by participants contained two values - positional error magnitude (in degrees), and direction. These values were sonically presented at every instantaneous point of the movement trajectory, whether the participant's hand was moving or not (note: against Effenberg's recommendations, as quoted above). As such, auditory feedback was effectively atemporal - complete error and directional values could in principle be extracted from a sample taken at any given instant. Unlike ecological information, understanding and use of these variables is not a product of the participant's continuous involvement in a perception-action task. Instead, comprehension of the message requires performance of a conscious, intellectual subtask. The participant must remember the target tone, tell whether the currently-heard tone is different (and how different), and decide which direction to move based on the remembered directionality of the mapping. If the participant can perform this subtask, the content of the message is accessible and appropriate action can be taken. In effect there is an intellectual barrier between perception and action, which may not be conducive to the development of fluent motor task performance (see: Ingold, 2001). Instead the learner may be required to consciously 'sample' the sound in a moment-to-moment fashion and update his/her motor performance using a strategy of discrete adjustments towards the target.

[30] There is an interesting discussion to be had about how classical literature on ‘knowledge of results’ and ‘knowledge of performance’ feedback may have influenced modern implementations of sonification for motor skill learning, and concurrent augmented feedback generally (see section 2.6.5).

Similar critiques of sonic information design have been made in the Auditory Display community in recent years. As an extreme example, recall Smith and Walker's (2002) sonified stock marked data (detailed in section 3.2.2). Making a decision based on the information heard requires performance of a substantial intellectual subtask.

Various researchers in the Auditory Display community have argued that this technique, termed "parameter-mapping sonification" (the transformation of measured parameters of the real world into sound) places too great a burden on the listener's auditory cognitive capacities, leading to dataset intelligibility problems for non-expert listeners (see: Barrass, 2012; Roddy & Furlong, 2014; Worrall, 2010; also section 3.2.3 in the current thesis). While the information required to make a decision is technically present, it is not easily accessible due to the requirement to remember and apply an abstract mapping rule.

7.2.6 The current experiment

The current experiment aims to test the hypothesised difference between a mapping designed to communicate the current state of the main performance metric and another designed for, broadly, the first-person perception/action challenges faced by an embodied agent. Previous experiments in this line of research have shown the benefits of attaching sound to movement in an otherwise silent motor task. Learning has been faster, better and possible to retain longer with the use of sound which corresponds more closely to the specific perception-action demands of the task, 4:3 bimanual rhythmic coordination (see Chapters 5 and 6).

In the previous experiment, performance was tested without feedback 24 hours later. It was observed that performance in the melodic sonification condition had declined substantially, to a level similar to control. However following playback of the target sound, good performance re-emerged in the melodic condition. The same benefit of listening was not observed in the rhythmic condition. Benefits in longer-term (post replay) retention for the more informationally-rich feedback support the idea that sound from a sonification is perceived in the context of the coordination pattern(s) it supports, and how it is relevant for task performance to an experienced listener. In other words, bodily skill is intrinsic to auditory perception. Experienced listeners, i.e. those who had learned how to use the given auditory information in an interactive task, could listen to the sound and perceive the movements of the task - effectively re-experiencing what it was like to perform the task correctly on the previous day (for a similar 'embodied' view of sound perception in movement sonification, see: Stienstra, Overbeeke, & Wensveen, 2011).

The advocacy of a perception-action approach to sonification is not intended to completely dismiss the common approach of ‘metric sonification'. The historical literature on the beneficial effect of KP and KR (mostly in the form of scores and verbal feedback) shows that there is value in this approach (Adams, 1987; Salmoni et al., 1984), and it is very plausible that an intellectual metric presented sonically could be helpful for learning. For example, a signal that the task is being performed well - even if the signal is not directly informative about movement quality - can indicate to the learner that whatever perception-action strategies he/she is currently utilising are effective. In this way, skill acquisition may yet proceed slightly faster than without sound.

However, it is my theory that in order to be *maximally* effective for the enhancement of motor performance and skill learning, sonification should become more than only another means to transmit knowledge for top-down implementation. The current experiment aims to test this. In one condition, the main measure of performance (bimanual ratio of movement timing) is sonified as a pure tone, and made available to moving participants as an auditory index of ongoing performance measurement – a kind of KP (hereafter referred to as the 'Metric' condition). In another condition, an event-coupled melody highlights the structure of task-intrinsic events which should be perceived and controlled in order to perform the task effectively without feedback[31] (hereafter referred to as the 'Melodic' condition). A third 'Control' condition is included in which participant movement is not sonified. Participants in

[31] Note also that the melody is technically redundant with respect to task-intrinsic event-related information. Based on the results of previous experiments in this series (the lack of an immediate guidance effect and the benefits of a replay for longer-term retention), one can be reasonably confident that this is useful information.

this condition hear constant, task-irrelevant pink noise. The inclusion of the control condition allows an assessment of the efficacy of either sonification method as feedback in its own right, regardless of the other.

Given the previous discussion, it is reasonable to hypothesise superior performance and learning (measured as lower bimanual ratio error at the end of the acquisition phase and in a subsequent no-feedback retention test) in the Melodic condition than in the Metric and Control conditions.

Following the removal of feedback, task performance is tested during the same session. Participants in the Melodic sonification condition should retain good performance, similar to previous experiments in this series. There are two possibilities for participants in the Metric condition following the removal of feedback. The first is that performance will remain largely similar. If participants make use of the tone in a manner similar to how a learner would make use of a coach's comments, or a score (i.e. as a metric, but not as perceptual information), then performance will not decline, as participants should continue to use intrinsic perceptual information to guide performance as they did during the practice phase. Removal of feedback should not interfere in this perception-action strategy. Park and Sternad (2015) found no evidence of a guidance effect in a similar task with numerical feedback. Their participants practiced a 3:2 bimanual coordination task and were given a number at the end of each trial corresponding to the mean bimanual ratio produced. They argue that the stable learning observed was a result of the primarily self-guided practice which made use of task-intrinsic perceptual information. The metric sonification in the current experiment can be considered a concurrent version of this kind of feedback.

The much-less-likely second possibility is a guidance effect. If participants use the tone as a source of perceptual information for the control of movement, performance should decline when it is no longer available, as there is no isomorphic feature of the task-intrinsic perceptual-motor workspace which might be coordinated with in the absence of augmented feedback. Examples of this result can be seen in studies which use Lissajous feedback (Kovacs et al., 2009; Maslovat et al., 2009).

Similar to previous experiments in this series, participants are tested again after a 24-hour delay. Given the differences in information structure between the two sonification systems discussed earlier (Melodic sonification as event-specifying perceptual information and Metric sonification as descriptive, less-specifying knowledge-of-performance), it is hypothesised that a replay of the target sound will improve performance in the Melodic condition but not at all the Metric condition. The sound of perfect performance, as presented via the constant tone, should not be perceivable as a guide for action any more than might a 350-yard drive report delivered prior to a golf swing.

7.3 Method

Design

The current experiment is composed of three independent groups. Participants in each of these groups undergo the same training and testing regimes, identical except for the sound information available during practice trials and prior to the later replay retention test. The three conditions are Melodic Sonification – in which correct performance produces a distinctive melody, Metric Sonification – in which correct performance produces a target tone of a recognisable pitch, and a Control condition – in which only pink noise is audible.

Participants

Sixty participants were recruited from the School of Psychology undergraduate population and from personal contacts of the experimenter. Undergraduate participants received partial course credit for their participation. 20 participants were assigned to each experimental condition on a pseudo-random basis. All participants recruited were right-handed. Handedness of participants was confirmed using a reduced version of the Edinburgh handedness inventory. Handedness scores were submitted to a one-way ANOVA. There was no effect of condition on scores: $F(2, 59) = .452$, $p = .639$. From this, it is possible to infer that handedness did not differ significantly between groups.

Potential participants who were drummers, professional musicians or dancers were excluded at the recruitment stage due to the influence these existing skills may have on performance in the experimental task. After the completion of testing, participants were administered a questionnaire on their experience of playing musical instruments and any extended dance practice. In the Melodic Sonification condition, 8 participants reported some musical experience (mean 7.63 years, SD = 5.76). Of these, 3 were current, active players (i.e. were engaged in playing their instrument(s) regularly), the rest having ceased active play an average of 8.5 years ago. In the Metric Sonification condition, 6 participants reported some musical experience (mean 4.50 years, SD = 2.65). Of these, 1 was an active player, the rest having ceased active play an average of 10.22 years ago. In the Control condition, 8 participants reported some musical experience (mean 7.38 years, SD = 6.61). Of these, 1 was an active player, the rest having ceased active play an average of 8.14 years ago.

Informed consent was obtained from all individual participants included in the study. Ethical approval for this study was granted by the School of Psychology Ethics Board at Queen’s University, Belfast.

Task and Apparatus

That task and apparatus were identical to those used in previous experiments (see Chapters 5 & 6). Reported here are alterations specific to the current experiment.

Feedback

Post-trial feedback was provided to all participants throughout the practice stage in the form of a graph of performance relative to perfect performance, as described in previous experiments.

Sonification in the Melodic Sonification condition was identical to the iteration described in the previous experimental chapter (see also Appendix A).

In the Metric Sonification condition, the performance measure, bimanual relative timing/frequency was sonified using a pure tone oscillator. As in the Melodic condition, perfect performance had a characteristic sound (a steady 440 Hz tone), which was played during each presentation of the demo and could be produced by perfect performance. During practice, the tone heard by participants was continuous, fluctuating according to the measured bimanual ratio. The activity of the tone over the time course of a practice trial reflected the form of the graph seen at the end of the same trial. Potential ratio values between 0.00 and 3.00 were scaled between 745 and 60 Hz; the perfect 4:3 ratio (1.33) produced 440 Hz. A ‘good’ trial would thus be accompanied by the sound of a tone wavering only very slightly around 440 Hz, whereas a ‘poor’ trial might be accompanied by wildly fluctuating pitches. To aid comprehension of the sonification, the performance measure (bimanual timing ratio) was explained to participants and the isomorphism between sound and graph highlighted.

Participants in the control condition heard only constant pink noise, at a comfortable volume, during demo presentation and practice. This was intended to mask the sound of the hands/fingers in physical contact with the apparatus, which could constitute task-relevant auditory feedback.

Procedure

As in previous experiments in this series, each participant followed an identical procedure which is outlined below in phases (see Figure 7.1). The only difference between experimental conditions was in the kind of sound audible during demo presentation, practice phase, and 'replay retention' phase.

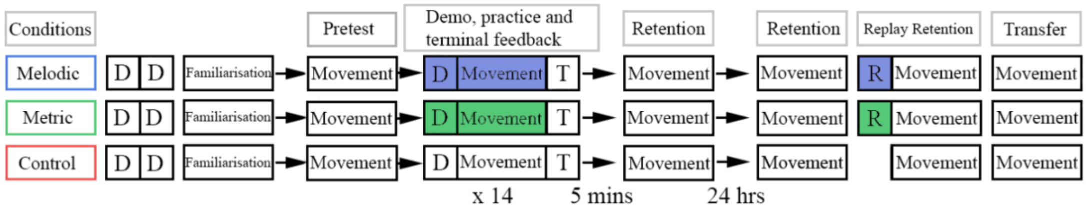

Figure 7.1: Schematic of the experimental procedure. D represents a play of the demo animation. T represents the graph (terminal) feedback. R represents a replay of the demo sound only. Shaded boxes indicate the presence of condition-appropriate sound/sonification. All unshaded 'Movement' boxes represent a practice/test trial conducted in the presence of only pink noise.

Analysis

Raw bimanual ratios (the primary measure of performance) were extracted as described in previous experiments for analysis. The absolute difference between each raw ratio and the ideal (4:3) ratio was calculated, then all differences for the trial were averaged to produce a representative single-trial score. In this arrangement, a lower bimanual ratio error score is desirable.

Mixed ANOVA are employed to test for effects of experimental condition assignment between testing occasions. One-way ANOVA are used to target effects of condition on trials of particular importance (practice trials 1 and 14). Where necessary, *t*-tests are employed for post-hoc decomposition of significant between-group effects. Direct comparisons of intergroup performance are made using *t*-tests at the final practice trial. Effect sizes are estimated using Eta squared for main effects and Cohen's *d* for simple effects. Rates of learning are compared by examining the slope values from curves fitted to performance data from individual participants during the practice stage.

The lasting effects of feedback through retention trials are identified with the use of a confidence interval-based statistical test of non-inferiority. This procedure makes it possible to assert (with 95% confidence) whether performance has *not* declined relative to the end of the practice phase, trial 14 (for a more detailed explanation of this procedure, see: Chapter 5; Walker & Nowacki, 2011). In this case, the non-inferiority interval for the Melodic

Sonification condition is set at .081, which is .5 times the difference in mean scores between the Melodic Sonification condition and Control at trial 14. The non-inferiority interval for the Metric Sonification condition is set at .034, which is .5 times the difference in mean scores between the Metric Sonification condition and Control at trial 14. If the upper confidence interval of the mean of difference scores between trial 14 and a later retention test falls below the non-inferiority interval value for the respective condition, then it is possible to infer that performance did not decline in the intervening time (i.e. inferring the absence of a guidance effect of feedback).

Some participants produced index finger marker movements which did not remain within the predefined trajectories between corner zones, leading to corner arrivals which were unidentifiable by metric extraction algorithms. The following participants' data were fully excluded before analysis and not replaced: One participant in the Control condition had his/her data eliminated on the transfer test. One participant in the Melodic condition had his/her data for trial 14 eliminated. One participant in the Metric condition had data eliminated for trials 4, 9 and 14; another had all data eliminated.

7.4 Results

Results of the experiment are presented and analysed in order, according to the chronology of the procedure. Figure 7.2 plots mean bimanual ratio error scores for each group across all stages of the experiment.

7.4.1 Pre-test

A one-way ANOVA performed only on data from the pretest trial revealed a significant effect of condition on error scores $F(2, 58) = 3.957$, $p = .025$. This indicates the presence of an unexpected difference between feedback conditions before practice. Post-hoc testing was administered ($\alpha = .016$; Bonferroni correction for three comparisons), which showed that the difference driving this effect was significantly worse performance in the Control condition (M = .485, SD = .141) relative to Metric Sonification (M = .356, SD =

.142), p =.009). No significant difference was observed between Control and Melodic Sonification (M = .388, SD = .163): p = .046, nor between Metric and Melodic Sonification conditions: p = .504.

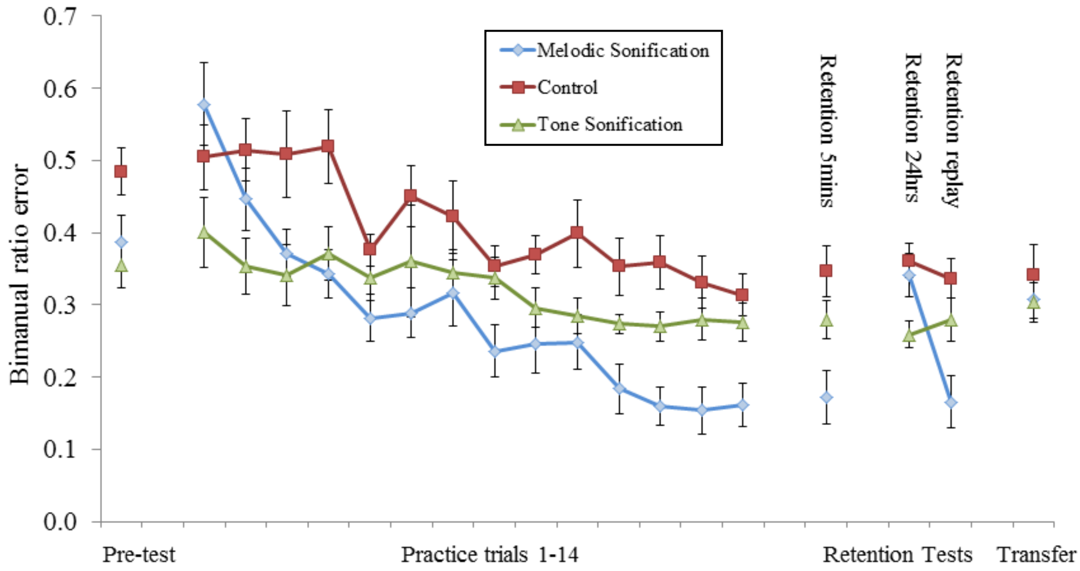

Figure 7.2: Rates of average bimanual ratio error for each experimental condition over pre-test, practice trials, retention testing and transfer. Error bars are standard error.

7.4.2 Practice trials 1-14

A mixed ANOVA with feedback condition as a between-subjects factor and trial number as a within-subjects factor was performed on acquisition data across all 14 trials. The assumption of spericity was violated, therefore Greenhouse-Geisser corrected values are reported where appropriate. ANOVA detected a significant main effect of trial: F(6.776, 352.348) = 18.164, p< .001, η^2 = .134; group: F(2, 52) = 7.883, p = .001, η^2 = .097; and a significant interaction between trial and group: F(13.552, 352.348) = 3.211, p< .001, η^2 = .049. Planned comparisons of intergroup performance at trial 14 were conducted to better understand the differental effects of feedback after practice. Alpha was set at .016 (Bonferroni correction for three comparisons). Comparisons indicated a significant difference between scores in the Melodic Sonification (M = .161, SD = .132) and Control (M = .314, SD = .129) conditions: t(37) = -3.67, p = .001, Cohen's d = 1.17; and between the

Melodic Sonification and Metric Sonification (M = .267, SD = .112): $t(35) = -2.64$, $p = .012$, Cohen's $d = .87$; but not between the Metric Sonification and Control conditions: $t(36) = 1.20$, $p = .238$. Together, these results show that error scores in the Melodic Sonification condition were lower than in the Metric Sonification and Control conditions by the end of the practice phase.

The time course of improvement in performance appears to vary between feedback conditions (see Figure 7.2), therefore curves were fitted to performance data from indiviual participants to determine average slopes for each condition. This curve-fitting procedure was performed iteratively as described in section 4.4.4. Holm-Bonferroni-corrected *p* values are reported for this analysis. In the Melodic Sonification condition, the mean slope value was .359, (S.D. = .787). A one-sided *t*-test revealed that slopes were significantly greater than 0: $t(19) = 2.042$, $p = .040$. In the Metric Sonification condition, the mean slope value was 2.468 (S.D. = 4.407), which was not significantly greater than 0: $t(19) = 2.505$, $p = .083$. In the Control condition, the mean slope value was 1.811 (S.D. = 4.041), which was not significantly greater than 0: $t(19) = 2.004$, $p = .083$. A one-way ANOVA revealed no effect of condition on slope values: $F(2, 59) = 1.921$, $p = .156$.

7.4.3 Retention testing

To test for the absence of a guidance effect in the sonification conditions, a statistical test of non-inferiority is performed, which compares the performance boost (relative to control) when sonification was available (at trial 14) to the same group's residual boost when sonification was withdrawn. The first set of tests concerns the short-term (5-minute delayed) retention test. In the Melodic Sonification condition, the mean of the difference scores between trial 14 and 5-min retention was .02, with 90% confidence intervals of [-.0002, .039]. The upper confidence interval (CI) here falls below the pre-set non-inferiority interval for the Melodic condition (.081); therefore performance in the Melodic condition statistically was not worse on this test than at trial 14. A significance test for this procedure can be performed via a one-sided, one-sample *t*-test on difference scores (between trial 14 and the test) relative to the non-inferiority interval (.081): $t(19) = .011$, $p = .011$. The same procedure is performed on the difference scores between trial 14 and 5-min retention in the Metric Sonification condition, yielding a mean difference score of .02 and 90% CIs of [-.0004, .040]. The upper CI of .040 falls above the non-inferiority interval of .034, which means that it cannot be said statistically that 5-min retention performance did not decline relative to trial 14 in the Metric Sonification condition ($t(19) = -.435$, $p = .668$).

On the 24-hour retention test one-way ANOVA revealed a significant main effect of condition on scores: $F(2, 58) = 4.851$, $p = .011$, $\eta^2 = .148$. Post-hoc analyses ($\alpha = .016$; Bonferroni correction for three comparisons) indicate that this effect is primarily driven by significantly lower error scores in the Metric Sonification condition (M = .259, SD = .081) relative to Control (M = .361, SD = .103): $p = .005$, Cohen's $d = 1.102$. The difference between Melodic Sonification (M = .342, SD = .133) and Control was not significant ($p = .581$); neither was the difference between Melodic Sonification and Metric Sonification ($p = .020$). The advantage observed in the Melodic Sonification condition relative to Control and Metric Sonification on the previous day was not evident on this test (M = .342, SD = .133).

On the replay-retention test, a significant main effect of condition was detected $F(2, 58) = 7.970$, $p = .001$, $\eta^2 = .221$. At this stage, error scores in the Melodic Sonification condition (M = .166, SD = .160) were significantly lower than those in the Metric Sonification condition (M = .279, SD = .131): $p = .013$, and the Control condition (M = .337, SD = .120): $p < .001$. The difference between Control and Metric Sonification was not statistically significant: $p = .199$ ($\alpha = .016$). To test the reactivated benefit of Melodic Sonification against the strength of that benefit on trial 14, 24 hours previously, a further test of non-inferiority was performed. The mean of difference scores was .013, with 90% CIs [-.014, .040]. The upper CI falls below the non-inferiority interval (.081), therefore it can be inferred that performance by participants in the Melodic condition after a sonic replay was not inferior to performance at the end of the sonified practice session 24 hours prior ($t(19) = -2.718$, $p = .014$. To test for a potential practice effect on error scores during the retention testing on day two, a paired-samples *t*-test was conducted using Control data from both tests. No difference in scores was evident between test 1 (M = .361, SD = .103) and test 2 (M = .337, SD = .120): $t(19) = 1.234$, $p = .232$.

No effect of feedback condition was detected on the transfer test: $F(2, 57) = .376$, $p = .688$.

7.5 Discussion

In the present experiment, melodic sonification coupled directly to movement events was more efficacious for performance enhancement in a bimanual task than a sonification mapping which presented the primary measure of performance through the medium of a tone.

7.5.1 Using the sonified metric

A major question around the Metric Sonification condition in the current experiment is whether participants would be able to understand and make use of the information about their performance as presented to them. The perception-action approach to sonification proposed here led to the prediction that, in general terms, performance in the Metric Sonification condition would be poorer than the Melodic condition, as a consequence of the sonic information being designed to communicate an abstract performance metric and therefore difficult to use for online control of action. The results from practice trials 1-14 support this prediction; participants in the Metric Sonification condition displayed significantly higher rates of error than those in the Melodic Sonification condition by the final trial.

The tone in the Metric condition was uncoupled from events in the intrinsic perceptual-motor workspace, therefore making its use independent of behavioural (i.e. perception-action) strategies for task performance (Newell et al., 1991). It was proposed instead that the Metric Sonification might be used similarly to a coach's comments (e.g. "good", "not good") or a rating, i.e. as an abstract, intellectual criterion by which performance could be judged (Adams, 1971). Using a similar task, Park and Sternad (2015) observed no dependence on similar feedback delivered terminally (as a single mean value rather than raw values delivered live as in the current experiment). From Figure 4, it can be seen that participants in the Metric condition displayed lower error rates on average than those in the Control condition on every practice trial. It might be possible therefore to infer that sonification of a performance metric was of some use to participants, relative to no useful sound at all. However this comparison is muddied by the unexpected finding of significantly lower pre-test error scores by participants in the Metric Sonification condition relative to the Control condition. This finding is difficult to explain, given the pseudorandom assignment of participants to each experimental condition and the similar levels of musical experience reported in both conditions. It is more likely that this group difference at pre-test reflects particularly poor performance by Control participants than a pre-existing advantage for participants in the Metric condition; consider that pre-test performance in the Control condition was also significantly worse than the Melodic condition, given an uncorrected alpha (p = .046). Metric and Melodic sonification mean scores did not differ significantly at pre-test.

The proposed efficacy of melodic, musical sonification of task-intrinsic events in bimanual coordination is lent further support by the results of the current experiment. Participants in the Melodic Sonification condition displayed significantly lower rates of error by the end of practice than in the Control condition. A test of statistical non-inferiority showed that performance in the Melodic condition did not decline in the absence of sonified feedback - further evidence against a short-term guidance effect with the use of this kind of feedback, replicating previous work in this series. Sonification of task-relevant events which would already exist in the task-intrinsic perceptual-motor workspace should make these events more salient, and the use of a melody should implicitly structure them in a way which is intuitively understood by learners (For other successful examples of musical sonification, see: Y. Chen et al., 2006; Scholz et al., 2014). Melodic sonification here encouraged the development of a perception-action strategy which was effective both when feedback was present, and when it was not, indicating that the desired isomorphism between sonic and task-intrinsic informational parameters was achieved (see also: Ronsse et al., 2011; van Vugt & Tillmann, 2015).

7.5.2 Metrics and music in retention

The finding of a lack of statistical equivalence between scores at early retention testing and trial 14 in the Metric Sonification condition may be void, due to the small size of the initial (statistically non-significant) difference between Metric Sonification and Control conditions at Trial 14. Conceptually, the test of non-inferiority requires some baseline level of effectiveness of a treatment relative to control – from which the non-inferiority interval is derived. The lack of such in this case may have set up a non-inferiority interval too small for the upper CI of the difference scores to fall beneath. This procedure was necessary to test for the absence of a guidance effect, and its failure here does not necessarily imply the opposite (i.e. the *presence* of a guidance effect in the Metric condition), especially considering the close similarity in mean difference scores between both sonification conditions and their associated CIs (M = .02 in the Melodic condition, with 90% confidence intervals of [-.0002, .039]; M = .02 in the Metric condition, with 90% CIs of [-.0004, .040]). A more likely explanation is that performance did not decline after the removal of Metric Sonification, however there was little relative benefit of sonification to begin with.

After a 24-hour delay, performance in the Melodic condition worsened substantially, reaching a rate of error comparable to the Control condition. As in previous experiments, participants frequently reported being unable to remember the melody which had assisted them on the previous day. The persistent benefit of melodically-sonified movement may be time-limited, or perhaps tied to task contexts in which melodic information is available to structure performance (to be listened to, not necessarily as feedback). Performance in the Metric condition did not exhibit a similar decline, instead remaining stable, with significantly lower error scores than the Control condition at this point. This may reflect a kind of learning which, while not displaying the same level of accuracy as the Melodic condition, is independent of feedback and also stable over the longer term (Park & Sternad, 2015). This fits with the 'coach's comments' interpretation of Metric Sonification (on this task, at least). A positive evaluation of performance can symbolically signal to the learner that whatever perception-action strategy they are using is effective, leading to its repetition (Adams, 1971, 1987). Although all participants, regardless of experimental condition, had access to such a signal in the post-trial graphical feedback, there may be some additional advantage to having the same available during movement.

After a replay of the sound of perfect task performance (6 cycles of the shapes over 18 seconds), good performance re-emerged in the Melodic condition. This was expected, and is a replication of a similar finding in previous work in this series. The extent of improvement was substantial; a test of non-inferiority showed that performance was no worse than when live sonification was available at the end of practice on the previous day. Lahav et al. (2013) demonstrated a similar benefit in piano learning; the current results show that the beneficial effects for motor retention of listening to a previously-learned musical piece can generalise to other complex tasks, given the appropriate couplings between real-world events and artificial sound.

The lack of improvement in the Metric Sonification condition after listening to the sound of perfect performance could fit with the notion that the tone was being used not in the style of ecological information (which would specify action patterns, thereby leading to a refresh), but perhaps as something more like a score, or intellectual metric (Park & Sternad, 2015). The powerful effect of listening to the sound of good performance in the Melodic condition indicates that perception of sound can serve action - where the listener is skilled enough to perceive a movement pattern from sonic information. In contrast, there was no statistical difference between performance in the Metric Sonification condition and Control after a replay. This lack of difference was not due to a practice effect which boosted performance in the control condition, as confirmed by a paired-samples *t*-test on data from both 24-hr tests in the Control condition which did not reach significance. Indeed, the mean error score in the Metric condition actually increased slightly after the replay of perfect performance (M = .259; .279 on test 24-hr tests 1 and 2 respectively). The lack of any improvement whatsoever in the Metric condition in contrast with the Melodic condition demonstrates the degree of separation that can exist between an experimenter/coach's conception of task performance and that of the perceiving-acting agent who performs it. The variables produced by measurement of task performance, although derived from real, physical activity, are not necessarily the same as informational variables picked up and used to coordinate action by a learner. As in a previous experiment in this series, improved performance in the Melodic condition did not transfer to a mirrored version of the task; there was no effect of feedback condition on scores.

7.5.3 Limitations of the current study

The current experiment has several limitations. Firstly, the choice to sonify in a manner consistent with ecological informational structure may not be as straightforward as is presented here. In the current, reduced, laboratory-friendly task, it suffices to mimic the temporal and structural characteristics of visual, proprioceptive and haptic information which is intrinsically and interactively available by performing the task and attending to the hands. However in more complex, elaborate real-world tasks, the informational variables which are of most use may be correspondingly elaborate, and perhaps a higher-order, multidimensional variable whose identity may not be immediately obvious, for example, *tau* (Lee, 1976), or the relative direction between two moving bodies (Wilson, Snapp-Childs, & Bingham, 2010) might be more appropriate for sonification in another task. Careful consideration and analysis of the task from a first-person perspective may be necessary to identify the appropriate variable(s) (see: Wilson & Golonka, 2013). Similarly, caution should be used when applying the results presented here to the sonification of performance metrics more generally. It is possible that in another context (using another mapping for presentation, or in a different task), metric information might be more readily interpretable and thus, more useful than the metric sonification described here.

It might be justifiably argued that continuous sonic information is less likely to be useful in a primarily rhythmic task like the one used in the current experiment due to the commonly reported benefits to sensorimotor timing performance afforded by the availability of discrete rhythmic information (Repp, 2005). However previous work in this series demonstrated (with the use of a rhythmic sonification which provided temporally-coupled sonic information with movements) that the coordination task used in the current experiment is different to and perhaps more complex than the kinds of tasks typically utilised in uni/bimanual tapping paradigms. The proposed benefit of pure action-sound coupling did not emerge relative to control. Wulf and Shea (2002) propose that sensory feedback strategies which are valid and useful in simpler experimental paradigms may not necessarily generalise to more complex tasks. The primary benefit of sonification in the current task may be unlocked through the use of a meaningful melodic structure, rather than through rhythmic/temporal information per se. The current experiment addresses a different, more obscure question relating to the level of description of the task (first-person or third-person) from which sonic information design should be conducted. The results provide some preliminary evidence in favour of the former, but further research is needed in more varied tasks and real-world skills to better support this recommendation.

7.5.4 Conclusion

Focussed investigations of different solutions for mapping and sonic information design are rare. Rarer still are comparisons motivated by competing, theory-driven hypotheses, which can explain findings and provide generalisable recommendations for mapping in other tasks. Intellectual understanding or knowledge-of-performance - i.e. how well one is performing a motor task relative to some abstracted criterion - can be a somewhat useful metric for a learner when presented via sonification. However more substantial benefits of sonification are available when the relationship between perception and action is catered to more directly; sonification which highlights task-intrinsic events and structures them in a meaningful way (e.g. with a melody) can engender performance of greater accuracy which is not dependent on the immediate presence of sound. If performance wanes after sonified training, it can be boosted again if the latter sonification strategy is used, but not the former. The use of a design strategy based in an embodied understanding of perception and action rather than internal cognition may be beneficial for varied fields which are interested in sonic information design. The current study indicates that further methodological validation by mapping comparison is necessary for sonification; had the Melodic condition not been included in the current experiment, it might be possible to conclude, based on performance in 24-hr retention, that sonification of a performance metric is a perfectly adequate strategy for augmented motor skill learning. Future work in this area should be cognizant of these issues.

Chapter 8

General Discussion

The preceding four chapters addressed the use of sonification in the learning of a novel motor task. In the current chapter, the results of these experiments will be discussed in light of the proposed perception-action approach to sonification for motor skill learning. To begin, the aims and findings of the investigations reported in the four empirical chapters are summarised. Then, broader questions about the use of sonification are addressed with reference to the results obtained. Finally, limitations of and suggested extensions to the current work are explored.

8.1 Empirical review

Chapter 4 reports an experiment in which a tabletop object-manipulation task was sonified, with the aim that sonification of the movements of the task would engender more accurate and lasting learning of the movements. Participants learned a sequence of movements with a set of plastic shapes, in a task designed to mimic the kinds of motor skills addressed by functional testing after stroke (Dobkin, 2004). In this instance, neither sonification of 1) shape arrivals nor 2) shape arrivals + movement trajectories improved motor performance relative to control. It was argued that the observed lack of performance enhancement in the sonification conditions was the outcome of a misguided conceptualisation of how sonification and augmented feedback generally aids motor performance. In short, this misconception entailed thinking about motor skill learning in terms of acquiring movement abilities, with less focus on how movement abilities are functionally underpinned by perception. Further analysis revealed that sonification of shape arrivals was a constraint on the spatial aspect of motor performance in this task which was not present in the control condition. In the control condition, participants need not have placed shapes as accurately to perceive correct performance. This difficult-to-accommodate constraint on action may have lowered scores on all primary performance measures. However, the movement style adopted in the presence of the sonic constraint (positional accuracy) was maintained into long term retention (up to 1 week later). In an extended discussion, an explanation of the task in perception-action terms was developed, and it was suggested that the notion of teaching 'movements' was misguided. Instead, participants may have learned to control sound. Going forward, bimanual coordination was identified as an ideal vehicle to carry on the study of sonification - through tasks in which motor performance is highly dependent on the structure of available perceptual information.

In chapter 5, a new task, 4:3 rhythmic bimanual shape-tracing, was implemented with sonification. This experiment had two primary aims: 1) to directly investigate the nature of the guidance effect in motor skill learning with sonification; 2) to test whether integration of the bimanual task demands into a single perceptual gestalt could alone account for any enhancement of motor performance shown in the sonification condition, i.e. to test whether live sonification was necessary in this task (Franz & McCormick, 2010; Franz et al., 2001). Participants learned to perform the task with significantly lower error scores when movements were sonified (a simple melodic pattern, when performed correctly) than control. The first aim was addressed through a five minute-delayed retention test, in which participants performed the task successfully without live feedback. This is clear evidence that a novel motor skill can be learned with sonification as feedback, and that learning is not dependent on the immediate presence of feedback - contrary to results observed with the use of certain forms of transformed visual information (Maslovat et al., 2009; Ronsse et al., 2011). The second aim was addressed by the inclusion of a third condition in which only the demonstration animation was sonified. Display of the sonified demo was intended to encourage perception of the motor task as a unified Gestalt, similar to the effect achieved by

Franz and McCormick (2010), who observed improved motor performance and less bimanual interference when the demands of a bimanual reaching task were expressed to participants in a 'unified' fashion. In the experiment reported here, unification through melodic sonification of the demo did not produce performance which was any better than that in the control condition. This confirmed that in the current task, there was a particular benefit of movements which produced sound. Interestingly, the benefit of live sonification in this experiment was time-limited. On a 24-hour retention test, performance in the sonification condition was indistinguishable from that in the control and sound-demo conditions. Participants reported that they were unable to remember the sound of good performance, and this was the main reason for their overnight decline.

Chapters 6 and 7 reported studies which extended the results of the previous experiment. These were aimed at understanding how exactly sonification might work to enhance motor performance. It was hypothesised that the structure of the information provided by sonification and its relation to the perceived goals of the task might be of particular importance.

The experiment reported in chapter 6 investigated whether there is a particular benefit of sonification which is conventionally musical - by comparing melodic sonification (a new melody, synthesised by physically-modelled strings) to purely rhythmic sonification (perceptually-indistinguishable white noise bursts). It was proposed that a simple coupling between movement events and sound events might enhance movement performance due to the superior temporal acuity of auditory perception relative to tactile/haptic perception. This effect has been harnessed in other sonified motor skills which have demonstrated improved performance when movements were sonified using a single type of sound (e.g. van Vugt & Tillmann, 2015). In this experiment, which used a more complex, multidimensional task, generic sonification of movement with bursts of white noise did not lead to performance improvements relative to control - the benefit of sonification was observed *only* in the melodic condition. This finding demonstrates that 'richness' of information in sonification mappings can be brought into use to serve motor performance, where there is a link between the structure of sonic information and the demands of the motor task at hand. Based on an embodied account of sound perception, it was hypothesised that a replay of the sound of good performance on day 2 might support successful movement coordination and effectively 'refresh' good performance in the sonification conditions. This strategy was successful in the melodic sonification condition, in which low-error performance (consistent with that at the end of practice on the previous day) did re-emerge after listening. This suggests that a sonification listener is an active perceiver, who can call upon embodied skill and prior physical practice to understand artificial sound and its relation to action - similar to how a skilled musician can get a sense of the movements of a performer from listening.

Chapter 7 reports the final experiment in the current thesis, which aimed to test competing theories of sonic information design. Sonification of a performance measure variable is a common strategy in the extant literature, showing mixed results (Rosati et al., 2012; Sigrist et al., 2013b). The distinction between sonification measure variables and sonification of task-intrinsic information/events is rarely considered in theoretical or empirical treatments of sonification for motor skill learning. It was proposed that the use of either sonification style would call upon on very different perception-action strategies. Although metric sonification might be available live, alongside movement, it could be used in a manner akin to more intellectualised forms of feedback like knowledge-of-performance: the signal contains a score to be extracted (Adams, 1987). In contrast, the Melodic sonification can be conceived as not having ‘content’ in itself; rather knowledge-of-performance is directly available in perception-action coupling with the system. In this experiment, the bimanual ratio produced by participants was sonified via the pitch of a sine tone, which was stable with perfect task performance and fluctuated with deviation from ideal task performance. This was compared to the melodic sonification used in previous experiments and a control condition. It was found that learning in the metric condition was not statistically superior to that in the control condition at the end of the practice stage, but learning was stable - error scores did not decline overnight and were significantly lower than control on day 2. Replay of the demo sound refreshed motor performance in the melodic condition, replicating a result of the previous experiment, but no benefit of sonic replay was observed in the metric condition. It was proposed that the two styles of sonification are reflective of different perspectives on task performance. Where understanding of melodic sonification was inextricable from the dynamics of the motor task, metric sonification represented an abstracted description of motor performance, as seen by an experimenter or coach. This characterisation is supported for the melodic sonification by the finding of a refresh on day 2 (for that condition), but is not as convincing in the case of metric sonification, due to the lack of a substantial benefit of sonification in the first place and lack of a decline to recover from.

8.2 Implications and recommendations for sonification mapping design in motor skill learning

Taken together, the results of these experiments provide support for the use of sonification as augmented feedback for motor skill learning. In chapters 6-8, learning of a novel skill was enhanced: lower error scores were observed with sonification of movement than without sound. Furthermore, good performance remained stable when sound was removed, overcoming the guidance effect, which was identified early in this research project as an issue in need of further investigation. The most interesting follow-up question from here is: *how* did mapping sound to movement aid performance? To answer this, it is proposed that one needs to consider the relevant constraints on perception and action. The following sections will provide broader interpretation of the results and synthesise some recommendations for sonification design.

8.2.1 The perils of metaphorical sonification

The relationship (mapping) between movement and sound in sonification for motor skill learning has not received the level of attention it arguably warrants in Psychology. The mapping can govern whether sonification is beneficial, has no effect, or is even a hindrance for performance. The first experiment in this thesis demonstrates some of the pitfalls of sonifying a new task to be learned. In designing the task and its associated sound, inspiration was taken from the existing literature on action-sound coupling and multimodal convergence in Psychology (e.g. Lahav, Saltzman, & Schlaug, 2007; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006), discussions in Auditory Display on the use of analogical mappings (e.g. Antle et al., 2009; Walker & Kramer, 2005) and the notion of cross-modal conceptual metaphors (Johnson, 2007; Leman, 2008). In Psychology, the results of perceptual judgement experiments and findings of sonic stimulus-response compatibility hint that certain sound-action pairings intuitively 'make sense'. For example, the 'SMARC effect' (Rusconi et al., 2006) suggests that there is some felt equivalence between a sound high in pitch and a physically high position. Burger, Thompson, Luck, Saarikallio and Toiviainen (2013) showed that similarly-embodied listeners perceive similar kinds of affordances for coordination with sound and music. Krueger (2014) sums up this concept:

"The acoustic structure of the music-as-heard thus determines the form of our musical advancing behavior; it shapes how we interact with and “inhabit” the music, experientially, and what we do with it." (p. 4)

This is not a controversial position; it is likely that most listeners would agree that it is difficult to march to a waltz (example from Krueger, 2014), and that it doesn't make sense to do yoga to EDM[32]. Certain sound morphologies seem to pair well in experience with certain kinds of movement; the two sometimes forming a single perceptual Gestalt (see section 2.7.5). I suggested that it should in principle be possible to exploit these links between sound and movement in sonification mapping design, to create sonification systems in which the interaction is intuitively understood. If the sound produced by the sonification system is in some way experientially congruent with the intended movement of the learner, then the result should be a coherent, unified interactive experience. While this argument seems plausible, and continues to guide research into perceived relations between movement and artificial sound (e.g. Frid, Bresin, Alborno, & Elblaus, 2016; Salgado-Montejo et al., 2016),

[32] Electronic Dance Music.

the metaphoric mapping approach as such may not be sufficient to support enhanced motor skill learning. To elaborate, a common explanation of the links between artificial sound and movement is that they are *conceptual*, i.e. that they exist primarily 'in the mind', as representations which get their meaning by being grounded in similar sensorimotor activity[33] (Antle et al., 2009; Johnson & Larson, 2003; Rusconi et al., 2006). This understanding of 'embodied-conceptual' sound perception dovetails with underlying theory in the motor learning literature, in which there can be a tendency to conceptualise skill as a stored motor program, and in which moving correctly is identical to motor learning[34] (see section 2.6.3). A focus on movement itself and *learning movements* as the marker of skill led to the search for sound morphologies which should make conceptual sense with the kinds of movements involved in the task. Hence, the approach behind sonification mapping decisions in Experiment 1 was ‘if conceptually-appropriate sounds can be paired with movement, then the movements of the task should be learned more effectively’.

The purpose of the prior retrospective on experiment 1 (section 4.6) was to explain as plainly as possible that conceptual links between sound and movement are not sufficient to guide sonic interaction design when the aim is enhancing skill acquisition. Following the lack of motor performance enhancement in the first experiment, and the finding that sound was an unexpected constraint on action, a re-examination of these assumptions was undertaken. This led to a renewed focus on perception and action processes in motor skill learning. Perhaps, a genuinely embodied approach to sonification should account for active bodily engagement as a constituent and necessary part of the task to-be-performed. For sound to be pragmatically useful for motor coordination, movement within the sonification system needs to generate perceptual information. Furthermore, this information needs to enable the learner to accomplish the goals of the task more easily (see also: Wilson & Golonka, 2015). Testing or intuiting what the goals of the task *are* for the learner - and the

---

[33] This is also a common understanding of embodied cognition generally, known in some circles as the "*conceptualisation hypothesis*" (see: Wilson & Golonka, 2013, p. 8).

[34] This is admittedly a simplified caricature of motor learning theory. It is employed here for expediency in the current discussion.

nature of the intrinsic information available which might enable him/her accomplish them - could be a good start for sonic interaction design based on this approach.

8.2.2 Sonification for musical skill

Experiments 2-4 present evidence that sonification can successfully enhance motor skill learning, but can also be taken as preliminary evidence in favour of the perception-action approach to design. Section 2.4 developed an explanation of motor skill learning as an 'education of attention' problem (E. J. Gibson, 1969). In other words, the challenge for a learner is, through exploration of the task system, to work towards the pickup and use of task-relevant information. In bimanual coordination tasks, the hands can be seen as performing two discrete tasks: tracing at different frequencies, moving in different directions, at different rates or out-of-phase etc. As shown by Mechsner, Kerzel, Knoblich and Prinz (2001), successful performance of these two tasks together is a product of being able to perceive the hands as coupled together - in a dynamic gestalt form. In the case of the experiment by Mechsner et al., this perceptual coupling was achieved by manipulating the form of the information available so that a complex bimanual relationship became a simpler, symmetrical coupling between two visible flags (actual hand movements were out of view). Symmetrical bimanual movements have long been known to be easier to control at high frequencies, as the visual information specifying symmetrical movement is easy to tune into (relative to anti-phase or any other kind of out-of-phase movement, see: Kelso, 1984). In this way, Mechsner et al. harnessed the known perception-action skills of their participants to make 1) the goals of the task and 2) control of movement perceivable and understandable. In related research, the same effect has been achieved by consolidating bimanual movement into a unified visual display which traces a recognisable shape (e.g. Kennedy, Wang, Panzer, & Shea, 2016). Without some transformation to make bimanual coupling easier to perceive (and therefore, control), these tasks are extremely difficult to learn (Kelso, Scholz, & Schoner, 1986).

Like the spatial and conceptual manipulations tested by Franz and colleagues (Franz & McCormick, 2010; Franz et al., 2001; see section 4.6.2), sonification can make task-relevant information more easily perceivable and understandable for the learner *but without* transformation of the kinematics of the task - in this case, by transposing important task events into music. Understanding how this style of sonification might have worked in experiments 2-4 is likely to be instructive from a design perspective. In Chapter 3, a discussion of sonic aesthetics was undertaken. This led to the suggestion that existing listening skill and prior experience using sound in real-world tasks could be exploited in sonification design. In theory, this could make the learner an active agent in the construction of their own meaning via interaction with the sonification system (Supper, 2014). From the earliest years of life, most people learn to sing, dance and clap along with simple melodies, like '*Happy Birthday*', or '*Twinkle Twinkle, Little Star*'. Simple melodies like these were composed to map onto the bimanual shape-tracing task used in experiments 2-4. The use of melodies may therefore have constrained behaviour (in the form of both information pickup and action to generate said information) towards interaction with those features of the task which are felt to be important in the context of musical behaviour. Furthermore, musical sonification may have helpfully constrained motor performance by making 'correct' sound production contingent on the required pattern of movement.

This is likely a similar effect to that demonstrated by Franz and McCormick (2010), in that participants' experience of the task may have been altered to make the task easier. 4:3 bimanual shape-tracing is an unusual task, one with which participants were unlikely to have had any prior experience. Making the task musical likely gives practice a very different phenomenological quality. With music, the instantiation of a familiar performative setting may make particular aspects of the task stand out as meaningful for the learner (such as the timing and ordering of movements relative to each other) where they might otherwise be overlooked in the unconstrained confusion of picking up a completely new skill. Here, the musical-aesthetic quality of the interaction served to constrain behaviour in line with the goals of the task, leading to measured improvement in motor performance.

Despite the reworking of the shape-tracing task into a musical task, the underlying kinematics of the motor task were not transformed, and so participants *did* actually learn the motor task, i.e. there was no evidence of a guidance effect upon the removal of feedback. The decline in performance after 24 hours displayed by participants who had practiced a musical version of the task might be taken as evidence of a *delayed* guidance effect attributable to sonification. However, the re-emergence of good performance after a sonic replay (not actual feedback) indicates that learning is not dependent on the presence of feedback per se (as in classic demonstrations of the effect), but suggests that it might be dependent on a form of musical practice. That *practice* possibly needs to be in some way perceptually available for skilful motor performance to emerge (Rietveld & Kiverstein, 2014).

8.2.3 Knowledge, information and variable selection in sonification

The first-person treatment of musical sonification described in this section is also useful for interpretation of the 'metric sonification' tested in experiment 4. I would argue that taking a first-person perspective is essential for sonic interaction design, despite that fact that behavioural measurements indexing performance improvement are always taken from the detached third-person perspective of the experimenter (or outside the lab, maybe the coach or clinician). Metric sonification in the form of a tone, mapped to the main measure of performance (bimanual timing ratio) was designed to conceptually mimic the classic 'knowledge' style of feedback, delivered from this detached perspective (KP/KR. See: Salmoni, Schmidt, & Walter, 1984). This style of feedback has much in common with examples of 'error sonification' in motor skill learning (e.g. Schmitz & Bock, 2014; Sigrist et al., 2013), in which skill acquisition is conceptualised as the ability to bring measured motor variables in line with an externally-defined metric of quality. By feeding back knowledge about how learner movements compare to the ideal movement profile (through sound), the aim is that he/she will adjust motor output accordingly. There are two related issues to consider regarding this approach to sonification mapping (which can also be applied to concurrent augmented feedback generally).

8.2.4 Improper variable selection

When considering variable selection for sonification, it can be useful to recognise that the motor variable being tracked by the researcher as a measure of performance may not necessarily correspond to the most useful informational variable(s) for the perceiving learner. In other words, measurement and experience are not isomorphic. An example of where this approach has led to misleading results can be found in Schaffert and Mattes (2015), in which a row-boat's acceleration time-course was parameter-mapped to pitch. It was concluded that making boat acceleration more perceptually available led to improvements in overall boat speed, since rowers could better synchronise during the recovery (post-stroke) phase. However in a letter to the editor, Hill (2015) showed that a strategy of increasing boat speed (to the degree reported) by controlling acceleration is biomechanically impossible. Instead, he argued, it was very likely that participants in the study were controlling oar propulsion force. In this rowing case, the measure variable was not the one being controlled by participants, and performance was enhanced by other means (possibly increased physical effort when sonification was available). Consider also pointing and reaching tasks (Boyer et al., 2016; Oscari et al., 2012; Schmitz & Bock, 2014). The task as instantiated in such research is to track or reach for a target, while using whatever information is provided by the system to guide one's effector/pointer. The variable of interest for measurement in this task is the absolute positional difference between hand/pointer position and target position (error). In sonified versions of the task, it is typically mapped to the pitch of a continuous tone heard throughout the reach. However, it is not certain that instantaneous positional error is a relevant variable for a moving individual in an everyday context. Everyday pointing for example is a primarily visuomotor task, with a criterion for success often defined in social terms (J. M. Kennedy, 1985). It makes sense from the detached perspective of experimenter to measure positional error as an objective performance index, but perhaps another, possibly higher-order variable might be more useful for the learner as a perceiver (Runeson, 1977).

The same argument can be made for the metric sonification used in experiment 4, however in that case, the difference between the structure of task-intrinsic information (and the usage strategies it invites) and that of the metric fed back is more obvious. The sonic information (i.e. the actual perceived variation in audible pitch rather than the score sonified) in the Metric condition was action-relevant only by way of an artificial convention established by the experimenter, through the choice of measurement variable, and the mapping between it and sound. On one hand, the measure of performance chosen (bimanual ratio) reflects how performance on similar tasks is assessed in the existing literature (D. M. Kennedy et al., 2013; Kovacs et al., 2010). On the other hand, the metric of performance studied by the experimenter can be arbitrary and abstracted from the perception-action interplay which characterises everyday skill acquisition in the first-person. The only marginal success of this mapping strategy in experiment 4 is a clear demonstration of this. In other metric sonifications (Konttinen et al., 2004; Oscari et al., 2012; Rosati et al., 2012; Schmitz & Bock, 2014; Sigrist et al., 2013b), the relationship between task-intrinsic experience and the metric fed back is similarly artificial. Performance measurements are usually taken from an external frame of reference, using conventionally-defined, but perceptually arbitrary units (cm, degrees etc.). This means that in order to understand the sonified feedback, the learner must adhere to the same conventions of, and frame of reference as, the experimenter.

8.2.5 How learners understand sonification

The second issue is related to the kinds of information which might be fed back via sonification and how they might be used. Controlling movement with the guidance of a sonified metric is not the same as coordinating movement with event-structured ecological information. Metric sonification, as in experiment 4, primarily describes what is happening - sometimes relative to a criterion for success (as in error sonification). A description of motor performance, even if delivered live and alongside movement, cannot so easily act as a constraint on action in the same way as can ecological information[35]. As mentioned in the introduction to Chapter 7, metric sonification is information with 'content'; that is, the sound contains a message which must be decoded according to a remembered mapping rule. This makes using metric sonification something of an intellectual task, similar to that faced by listeners of parameter-mapped auditory displays (see section 3.2.2). The information provided by metric sonification may not be immediately obvious to the learner as the sound heard is secondary to the number(s) it is intended to convey. The possible consequences of this approach to sonification mapping in motor skill learning is slow (but possibly stable) learning, as demonstrated in experiment 4 by comparing metric sonification to melodic across practice and retention.

In contrast, more basic sonic information coupled to and structured by events, while not designed to transmit a movement parameter in itself, can enable finer control over the same movement parameter. In the melodic sonification conditions in experiments 2-4, participants were able to access knowledge about motor performance through interaction with the system. The benefit of melody relative to purely rhythmic sonification in experiment 3 is evidence of the importance of richness and specificity of information - the structure of information provided via sonification in these conditions was matched to the structure of the task. Differentiable patterns in sound (e.g. perceptually separable tones) should allow learners to differentiate specific and relevant events in the perceptual-motor workspace. In this way, sonic information can be meaningful through use and enable finer perception of motor performance which is also 'direct' (see section 2.4.3).

[35] This is not to say that metric sonification could *never* be used in this way; that is, in the way one might use 'information' in the Ecological sense. However, given that the feedback by its nature is abstract and its use can rely on intellectual decoding strategies, the learning process would likely be prohibitively slow, and a guidance effect likely.

8.2.6 Section summary

The perception-action approach of the current thesis has been in the service of proposing a particular kind of 'embodied' model for sonification design in the context of motor skill learning. This approach seeks to solve what is described in Auditory Display as the 'mapping problem' - the inability of the field to come up with broadly-accepted conventions in regard to mapping. The current thesis will not solve the mapping problem for Psychology, but by highlighting important (sometimes overlooked) theoretical and methodological issues in the field, it might guide sonification designers towards the creation of more effective prototypes for motor skill learning enhancement.

In this section I have argued that the proposed theoretical framework provides a convincing basis for explaining how sonification has enhanced (or not enhanced) motor performance in the reported research and guides learner experience, while making some general recommendations for design. The degree of 'embodiment' in the approach advocated here is relatively strong, as compared to some other circles which use the term in reference to conceptual metaphor theory (Johnson & Larson, 2003; cf. Wilson & Golonka, 2015). This is important to stress here as the study reported in chapter 4 suggests (tentatively) that conceptual metaphor in sonification cannot stand in for couplings between perception and action in interaction with the environment. If motor skill learning is a perception-action phenomenon, then it follows that sonification should provide information structured by the task-intrinsic events which would need to be perceived if the task were learned without feedback. A clearer definition of what 'information' is (according to the Ecological approach) can help guide the design of sonified feedback whereby knowledge is a product of interaction rather than transmission. As a final recommendation, it should be acknowledged that learners have a wealth of socioculturally-situated listening experience already; it can therefore be beneficial to incorporate familiar musical styles into sonified feedback.

8.3 Limitations and possible extensions of the research

This section will address some limitations of the research reported in this thesis. Where possible, further studies will be suggested to shore up the approach.

8.3.1 Scope

The failure of the first study (reported in Chapter 4) to demonstrate a benefit of sonification was responsible in large part for the structure of this thesis, for good and ill. On the one hand, it led to the formulation of what I see as a stronger theoretical model of sonification than might have existed otherwise. On the other, this early failure limited the scope of the subsequent experiments. It is unlikely that learning in bimanual coordination tasks is a special kind of motor skill learning, somehow different psychologically to learning in less-constrained, 'real-world' motor skills. Complex bimanual coordination is merely an example of a task in which the information which is critical for motor coordination is particularly difficult to perceive. As sonification was utilised to make that information more salient, or easier to perceive in the shape-tracing task, so too can it be utilised in a wider range of motor skills with varied perception-action requirements[36]. However, the experiments which would further validate this approach (and confirm that it 'scales up') in sporting or rehabilitative contexts remain to be done. The interested reader could consider the ongoing research from the labs of Danna et al. and Effenberg et al., whose approaches to sound design in larger-scale motor tasks (writing; rowing respectively) would fit with the approach advocated here in some ways (e.g. Danna et al., 2015; Effenberg, Fehse, Schmitz, Krueger, & Mechling, 2016).

8.3.2 Situated skill in the use of sonification

A limitation of the current set of studies is that there are still some questions unanswered surrounding the use of sound morphologies which might be particularly useful for guiding attention for those attuned to a certain cultural/situational context. Certainly, musical/melodic listening is an example of a skill or cultural practice which was harnessed

[36] However, the design challenge will likely be greater.

in service of learning a novel task in the experiments reported in chapters 5-7, but it is difficult to make a purely aesthetic argument in favour of a particular kind of sonification design from these results alone. In experiment 3, which pitted musical sonification of movement against purely rhythmic sonification, it is likely that the complementary relation between the structure of the task and that of the sonic information provided was a major factor in the enhancement of motor performance – rather than the engagement of a culturally-situated musical listening skill per se, although it likely contributed to the ease of use of the system. There could be an experiment designed which would address this potential for sonification-task design directly, in which aesthetics and informational specificity were not as confounded. However, the extent to which it is possible to separate these two concepts in human experience is highly debatable, given that the current thesis takes something like a pragmatist view of aesthetics as inseparable from use (Dewey, 1934). A solution to this might be possible with the recruitment of a group of domain-specific experts and the use of sound morphologies which cater to their specific subculture. If the experts show a distinct advantage relative to non-experts on a mutually novel motor task, then a stronger case could be made. There is a reasonable expectation that an effect of skill *would* be found, if we consider some contemporary research in auditory perception and action (Cesari et al., 2014; Su & Pöppel, 2012).

An extension of this research which could further ground understanding of sonification in musical practice could address the 'refreshing' effect of a replay in retention. In section 8.2.2, I suggested that a decline in performance after 24 hours (and re-emergence after a sonic replay) was not indicative of the guidance effect in the traditional sense; the perceptual availability of a known musical practice could be a requirement for the emergence of skilful behaviour, rather than the presence of feedback. To explore this interpretation further, an experiment could be conducted in which some kind of broader musical practice was established in the acquisition stage. It could then be tested whether good performance re-emerged after a short reintroduction to that context. The form taken by this practice would need to be considered very carefully in order that it might functionally support performance while also allowing the researcher to maintain some semblance of experimental control. As a simple example, a participant could learn the 4:3 shape-tracing task with sonification and a backing track - perhaps something as simple as a drum beating in 4:4 time, or some other dynamic structure with respect to which motor performance could be coordinated (this need not necessarily be sound)[37]. If task performance declines after 24 hours as before (tested under control conditions of no feedback and no backing track), the participant could subsequently be exposed to the version of the backing track which would be available alongside perfect task performance before being tested under control conditions for a second time. If performance improves in this context, then it should be possible to claim with greater certainty that skilful motor performance draws in part upon situated practices. This might also serve to deflect an alternative theoretical interpretation of the refresh effect in experiments 3 and 4 - that the sound of good performance simply reactivated a stored motor program. The refresh effect might be achievable without live feedback *and* without an actual replay of the direct consequences of correct performance (the sound of the demo). If so, then the results might form the basis of an updated conceptualisation of the guidance effect as a consequence of situated embodiment and an impoverished testing environment.

8.3.3 Moving with sound, but not sonification

It was not tested in the current thesis whether similar enhancements in performance could be achieved through moving along *with* structured auditory information. Most people find it very easy to synchronise sequenced, rhythmic behaviour with respect to an external auditory stimulus, whether a metronome or more structurally-rich music (Repp, 2005; Van Dyck et al., 2013). In the shape-tracing task, it would be useful to test whether the sound of the demo, played continuously throughout the movement phase in every trial, would better support task performance than control conditions – or perhaps, support performance to a

[37] The obvious solution might be to take a reductive approach and use a metronome. However there is some evidence that the rigid timing constraints imposed by metronomes can increase the coordinative difficulty of bimanual tasks; see Kovacs et al. (2010). Perhaps instead, the backing 'track' could respond dynamically in some way to the pace of the learner, the two thereby forming a higher-order, coupled system.

similar degree as sonification. It is possible however, that the complex nature of the task and the constraints of the task goals (movements in space, time, and in a particular order) might necessitate movement-coupled feedback – for perception of error alone. Furthermore, without the interactive experience of using live sensory feedback, how else might participants come to learn the specifying nature of the auditory demo information played during practice? Participants could face the same symbol-grounding problem lamented by sound designers in Auditory Display, as the ecological relations between sound and action structure would not be enforced (Roddy & Furlong, 2014).

8.3.4 Transfer effects

Although transfer of learning was not a central focus of the research reported here, research is needed to understand the conditions under which learned skills can transfer from one task to another when sonification is employed. The 'motor' interpretation of transfer employed in the current thesis was adopted uncritically from the classical motor skill learning literature, when perhaps a more interesting and useful version could be conceived in terms of education of attention. Snapp-Childs, Wilson and Bingham (2015) propose that successful transfer of learning from one motor task to another is a related to the presence of common perception-action dynamics in both the learned and unlearned tasks. Upon encountering the new task, an experienced performer should perceive familiar structures and relations between perception and action, which may speed up learning of the new task, but might not result in accomplished performance of the new task right away. For example, it is generally found that violinists can more quickly pick up and learn the cello than musical novices, but will not be able to play a familiar piece on the cello right away. What has transferred is a tacit knowledge of how one's behaviour should be constrained in the context of an interaction with a stringed instrument, rather than a program of knowledge for the performance of a piece (Ingold, 1996). The same understanding could be applied to the task used in the current research by allowing a series of practice trials on the mirrored transfer test and analysing the learning curve, rather than testing performance on a single trial. This is an alternative kind of transfer which would not show up in the kinds of transfer tasks utilised in the research presented here.

8.3.5 Applications outside the lab

The theoretical approach and empirical findings presented in the current thesis have application in elite sports training and movement rehabilitation. The former application is arguably more mature from a research perspective, now that some trials of movement sonification are being performed with the involvement of professional athletes (see for examples: Effenberg et al., 2016; Schaffert & Mattes, 2015; Stienstra et al., 2011).

As demonstrated by some of the published literature showing a lack of performance enhancement in sport with sonification (e.g. Sigrist, Fox, Riener, & Wolf, 2016; Sigrist et al., 2013b), there is still need for focussed research to establish clearer methodological practices in this area[38]. Taking a perception-action approach as outlined in this thesis may enable sporting professionals to tune their performance by listening to the sound they create in practice. Wearable or otherwise mobile sonification systems have been developed or are currently in development, which could facilitate an uptake in the technique (Baudry et al., 2006; Danna, Fontaine, et al., 2015; Effenberg, Schmitz, Baumann, Rosenhahn, & Kroeger, 2015; Horsak et al., 2016). Even if a sonification feedback system in itself were not mobile, the ‘refresh effect’ demonstrated in experiments 3 and 4 suggests that there could be value to listening to the sound of sonified performance before trying the task in the field. An .mp3 recording played through headphones could support good performance where the system itself might be too cumbersome.

Research continues to show benefits of sonification in motor rehabilitation. In stroke, broadly positive findings have been reported on the use of sonification as, in effect, a sensory substitution device to retrain upper limb function (Y. Chen et al., 2006; Maulucci &

[38] The fact that studies like these (which report null results on the effect of sonification) are published and available is extremely valuable; as they can still (and have done here) inform future design practices. There are likely many more such datasets languishing on hard drives due to the difficulty of publishing null results.

Eckhouse, 2001; Robertson et al., 2009; Rosati et al., 2013; Schmitz et al., 2014). Part of the potential benefit of sonification in this case could be the more engaging nature of sonified rehabilitation as compared to traditional practice (Scholz et al., 2014). Musical interaction in particular could be efficacious here; there is much unrealised scope for the use of music in movement sonification, through which could be designed richly-meaningful and engaging forms of rehabilitation (see section 3.3.7 for discussion).

In Parkinson's disease, sonification has shown promise as an aid for the coordination of gait. Traditionally, rhythmic auditory cueing (synchronisation of gait with a metronome) has been employed here, but recent research suggests that sonification could assist by providing richer, more action-relevant auditory structure (Rodger & Craig, 2016; Rodger et al., 2014). The further development of mobile systems could facilitate wider adoption here as well (e.g. Horsak et al., 2016).

8.4 Conclusion

The work presented in the current thesis is intended to begin the development of a theoretical and methodological framework for sonification mapping design in motor skill learning. The lack of such a framework, which could both constrain mapping choice and facilitate comparisons between mappings, has been identified by other authors as necessary for the field to mature, but currently lacking (Boyer, 2015).

A case has been made here that the learner's experience of the task is a vital consideration in mapping design. I propose that an embodied and aesthetic approach – which is inclusive of the perceiving-acting agent, the supportive information structures in the immediate environment and available cultures of practice – can allow designers to account for (something close to) experience. When I talk about experience, I do not refer to something entirely 'inner' and private to the skilful actor. The proposed approach states that a variety of distributed and embodied systems in the world are both constituent parts of - and necessary for - experience, which is itself seen as *enacted*, rather being than something that happens *to* us, or inside us (Noe, 2004).

The above considerations about the makeup of experience may not in principle be truly separable from the forms of movement produced by an expert when an inclusive analysis of skill is performed (Ingold, 2001). The purposeful activity which is implicated in perception and action with sonification (indeed, the same actions measured by behavioural scientists) is constrained by how the world shows up for the learner. The case has been made that the nature of the task, in combination with the skills and habits of the individual, is a compelling mechanism for the form of this activity (see section 2.5).

Sonification can be used to alter the task, by augmenting some informational structures to be more readily picked up and used by an agent with the appropriate skills (such as musical listening experience); in effect, guiding perception and action in line with the coordinative goals of the motor skill to be learned. With this approach, sonification can enhance motor skill performance without leaving learners dependent on feedback. Additionally, longer term skill retention can be facilitated through sonic playback. In the experimental literature, the majority of sonification prototypes reported use aesthetically impoverished or abstract sound types, and sometimes sonify inappropriate motor variables. There is therefore great scope for future application of this approach in research and practice.

Bibliography

Abernethy, B., Thomas, K. T., & Thomas, J. T. (1993). Strategies for Improving Understanding of Motor Expertise (or Mistakes We Have Made and Things We Have Learned!!). In J. L. Starkes & F. Allard (Eds.), *Cognitive Issues in Motor Expertise* (pp. 317–356). Elsevier Science Publishers B.V.

Adams, J. A. (1971). A Closed-Loop Theory of Motor Learning. *Journal of Motor Behavior*, *3*(2), 111–150. http://doi.org/10.1080/00222895.1971.10734898

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74. http://doi.org/10.1037/0033-2909.101.1.41

Adolph, K. E., & Kretch, K. S. (2015). Gibson's Theory of Perceptual Learning. In H. Keller (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., Vol. 10, pp. 127–134). New York, NY: Elsevier. http://doi.org/10.1016/B978-0-08-097086-8.23096-1

Agostini, T., Righi, G., Galmonte, A., & Bruno, P. (2004). The Relevance of Auditory Information in Optimizing Hammer Throwers Performance. In P. B. Pascolo (Ed.), *Biomechanics and Sports* (pp. 67–74). Vienna: Springer Vienna. http://doi.org/10.1007/978-3-7091-2760-5_9

Albright, T. D., Kandel, E. R., & Posner, M. I. (2000). Cognitive neuroscience. *Current Opinion in Neurobiology*, *10*(5), 612–624. http://doi.org/10.1016/S0959-4388(00)00132-X

Altenmüller, E. (2007). From the Neanderthal to the concert hall: Development of sensory motor skills and brain plasticity in music performance. In *Proceedings of the International Symposium on Performance Science. Utrecht: European Association of Conservatoires (AEC)* (pp. 1–14).

Anderson, D. I., Magill, R. A., Sekiya, H., & Ryan, G. (2005). Support for an explanation of the guidance effect in motor skill learning. *Journal of Motor Behavior*, *37*(3), 231–8. http://doi.org/10.3200/JMBR.37.3.231-238

Antle, A. N., Corness, G., Bakker, S., Droumeva, M., van den Hoven, E., & Bevans, A. (2009). Designing to support reasoned imagination through embodied metaphor. In *Proceeding of the seventh ACM conference on Creativity and cognition - C&C '09* (p. 275). New York, New York, USA: ACM Press. http://doi.org/10.1145/1640233.1640275

Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, *21*(6), 1399–1407. http://doi.org/10.1016/S0896-6273(00)80658-3

Bach-y-Rita, P., & Kercel, S. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, *7*(12), 541–546. http://doi.org/10.1016/j.tics.2003.10.013

Ballass, J. (1994). Delivery of Information Through Sound. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification and Auditory Interfaces* (pp. 79–94). Reading, MA: Addison-Wesley.

Ballora, M. (2016). Music of migration and phenology: Listening to counterpoints of Musk Ox and Caribou migrations, and cycles of plant growth. In *Proceedings of the 22nd International Conference on Auditory Display (ICAD-2016).*

Ballora, M., & Kenney, M. (2014). Changes in Antarctic Ice - 400,000 years to present Presented at PSU Polar Day 2014.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., … Altenmüller, E. O. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–926. http://doi.org/10.1016/j.neuroimage.2005.10.044

Barnett, M. L., Ross, D., Schmidt, R. A., & Todd, B. (1973). Motor skills learning and the specificity of training principle. *Research Quarterly*, *44*(4), 440–7.

Barrass, S. (2012). The aesthetic turn in sonification towards a social and cultural medium. *AI and Society*, *27*(2), 177–181. http://doi.org/10.1007/s00146-011-0335-5

Barrass, S., & Kramer, G. (1999). Using sonification. *Multimedia Systems*, *7*(1), 23–31. http://doi.org/10.1007/s005300050108

Barrass, S., & Vickers, P. (2011). Sonification Design and Aesthetics. In Hermann, A. Hunt, & J. Neuhoff (Eds.), *The Sonification Handbook* (1st ed.). Berlin: Logos Publishing House.

Baudry, L., Leroy, D., Thouvarecq, R., & Choller, D. (2006). Auditory concurrent feedback benefits on the circle performed in gymnastics. *Journal of Sports Sciences*, *24*(2), 149–

156. http://doi.org/10.1080/02640410500130979

Bavelier, D., Dye, M. W. G., & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences*, *10*(11), 512–518. http://doi.org/10.1016/j.tics.2006.09.006

Bellotti, V., Back, M., Edwards, W., Grinter, R., Lopes, C., & Henderson, A. (2002). Making sense of sensing systems: five questions for designers and researchers. In *Proceedings of ACM Conference on Human Factors in Computing Systems* (pp. 415–422).

Beniger, J. R., & Robyn, D. L. (1978). Quantitative Graphics in Statistics: A Brief History. *The American Statistician*, *32*(1), 1–11.

Bernstein, N. (1967). The co-ordination and regulation of movements. Oxford, NY: Pergamon Press.

Bilodeau, E., & Bilodeau, I. (1958). Variable frequency of knowledge of results and the learning of a simple skill. *Journal of Experimental Psychology*, *55*(4), 379–383.

Bilodeau, E., Bilodeau, I., & Schumsky, D. (1959). Some effects of introducing and withdrawing knowledge of results early and late in practice. *Journal of Experimental Psychology*, *58*(2), 142–144.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11818–11823. http://doi.org/10.1073/pnas.191355898

Bootsma, R. J., Fernandez, L., & Mottet, D. (2004). Behind fitts' law: Kinematic patterns in goal-directed movements. *International Journal of Human Computer Studies*, *61*(6), 811–821. http://doi.org/10.1016/j.ijhcs.2004.09.004

Boyer, É. O. (2015). *Continuous auditory feedback for sensorimotor learning*. L'Université Pierre et Marie Curie.

Boyer, É. O., Babayan, B. M., Bevilacqua, F., Noisternig, M., Warusfel, O., Roby-Brami, A., … Viaud-Delmon, I. (2013). From ear to hand: the role of the auditory-motor loop in pointing to an auditory source. *Frontiers in Computational Neuroscience*, *7*(April), 26. http://doi.org/10.3389/fncom.2013.00026

Boyer, É. O., Bevilacqua, F., Susini, P., & Hanneton, S. (2016). Investigating three types of continuous auditory feedback in visuo-manual tracking. *Experimental Brain Research*. http://doi.org/10.1007/s00221-016-4827-x

Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organization of sound*.

Cambridge, MA: The MIT Press.

Bregman, A. S. (2001). Auditory Scene Analysis. In N. J. Smelzer & P. B. Bates (Eds.), *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Pergamon Press. http://doi.org/10.1016/B0-08-043076-7/00663-X

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, *89*(2), 244–249. http://doi.org/10.1037/h0031163

Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, *8*(August), 599. http://doi.org/10.3389/fnhum.2014.00599

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4*, 183. http://doi.org/10.3389/fpsyg.2013.00183

Cadoz, C. (2009). Supra-Instrumental Interactions and Gestures. *Journal of New Music Research*, *38*(October 2012), 215–230. http://doi.org/10.1080/09298210903137641

Cage, J. (1996). *Musicage. Cage muses on words, art, music*. (J. Retallack, Ed.). Hanover, NH: Wesleyan University Press.

Carello, C., Wagman, J., & Turvey, M. T. (2005). Acoustic specification of object properties. In J. D. Anderson & B. Fisher Anderson (Eds.), *Moving image theory: Ecological considerations* (pp. 79–103). Southern Illinois University Press.

Cascone, K. (2004). Grain, Sequence, System: Three Levels of Reception in the Performance of Laptop Music. *Contemporary Music Review*, *22*(4), 101–104. http://doi.org/10.1080/0749446032000156900

Castaneda, B., & Gray, R. (2007). Effects of Focus of Attention on Baseball Batting Performance in Players of Differing Skill Levels. *Journal of Sport & Exercise Psychology*, *29*(1), 60–77.

Castiello, U., Giordano, B. L., Begliomini, C., Ansuini, C., & Grassi, M. (2010). When ears drive hands: The influence of contact sound on reaching to grasp. *PLoS ONE*, *5*(8), 1–9. http://doi.org/10.1371/journal.pone.0012240

Cecchetti, L., Kupers, R., Ptito, M., Pietrini, P., & Ricciardi, E. (2016). Are Supramodality and Cross-Modal Plasticity the Yin and Yang of Brain Development? From Blindness to Rehabilitation. *Frontiers in Systems Neuroscience*, *10*(November), 1–8. http://doi.org/10.3389/fnsys.2016.00089

Cesari, P., Camponogara, I., Papetti, S., Rocchesso, D., & Fontana, F. (2014). Might as well jump: Sound affects muscle activation in skateboarding. *PLoS ONE*, *9*(3). http://doi.org/10.1371/journal.pone.0090156

Chadabe, J. (2002). The limitations of mapping as a structural descriptive in electronic instruments. In *NIME '02 Proceedings of the 2002 conference on New interfaces for musical expression, May 24-26, Dublin, Ireland* (pp. 1–5).

Chamberlin, C. J., & Magill, R. A. (1992). A note on schema and exemplar approaches to motor skill representation in memory. *Journal of Motor Behavior*, *24*(2), 221–4. http://doi.org/10.1080/00222895.1992.9941617

Chang, J.-Y., Chang, G.-L., Chang Chien, C.-J., Chung, K.-C., & Hsu, A.-T. (2007). Effectiveness of two forms of feedback on training of a joint mobilization skill by using a joint translation simulator. *Physical Therapy*, *87*(4), 418–30. http://doi.org/10.2522/ptj.20060154

Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, *7413*(December). http://doi.org/10.1207/S15326969ECO1502

Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Chen, J. L., Rae, C., & Watkins, K. E. (2012). Learning to play a melody: An fMRI study examining the formation of auditory-motor associations. *NeuroImage*, *59*(2), 1200–1208. http://doi.org/10.1016/j.neuroimage.2011.08.012

Chen, Y., Huang, H., Xu, W., Wallis, R. I., Sundaram, H., Rikakis, T., … He, J. (2006). The design of a real-time, multimodal biofeedback system for stroke patient rehabilitation. In *Proceedings of the 14th annual ACM international conference on Multimedia - MULTIMEDIA '06* (p. 763). New York, New York, USA: ACM Press. http://doi.org/10.1145/1180639.1180804

Chiou, S. C., & Chang, E. C. (2016). Bimanual coordination learning with different augmented feedback modalities and information types. *PLoS ONE*, *11*(2), 1–19. http://doi.org/10.1371/journal.pone.0149221

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, *33*(March), 269–298. http://doi.org/10.1146/annurev.neuro.051508.135409

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, *3*(9), 345–351. http://doi.org/10.1016/S1364-6613(99)01361-3

Clark, A. (2003). *Natural Born Cyborgs: Minds, Technologies, and the Future of Human*

Intelligence. New York, NY: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(January), 7–19. http://doi.org/10.1111/j.1467-9744.2009.01021.x

Cook, P. (2001). Principles for designing computer music controllers. *Proceedings of the International Conference on New Interfaces for Musical Expression 2001 (NIME 2001)*, 1–4. http://doi.org/10.1101/gr.075622.107

Correia, V., Araújo, D., Cummins, A., & Craig, C. M. (2012). Perceiving and acting upon spaces in a VR rugby task: expertise effects in affordance detection and task achievement. *Journal of Sport & Exercise Psychology*, *34*(3), 305–21.

Culham, J. C., & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology*, *11*(2), 157–163. http://doi.org/10.1016/S0959-4388(00)00191-4

Danna, J., Fontaine, M., Paz-Villagrán, V., Gondre, C., Thoret, E., Aramaki, M., … Velay, J.-L. (2015). The effect of real-time auditory feedback on learning new characters. *Human Movement Science*, *43*, 216–228. http://doi.org/10.1016/j.humov.2014.12.002

Danna, J., Paz-Villagrán, V., Capel, A., Pétroz, C., Gondre, C., Pinto, S., … Velay, J.-L. (2014). Movement Sonification for the Diagnosis and the Rehabilitation of Graphomotor Disorders. In M. Aramaki, O. Derrien, R. Kronland-Martinet, & S. Ystad (Eds.), *Sound, Music, and Motion* (Vol. 8905, pp. 246–255). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-319-12976-1

Danna, J., Paz-Villagrán, V., Gondre, C., Aramaki, M., Kronland-Martinet, R., Ystad, S., & Velay, J.-L. (2015). "Let Me Hear Your Handwriting!" Evaluating the Movement Fluency from Its Sonification. *PloS One*, *10*(6), e0128388. http://doi.org/10.1371/journal.pone.0128388

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*(10), 460–467. http://doi.org/10.1016/j.tics.2003.08.014

Dewey, J. (1934). *Art as experience*. New York, NY: Minton, Balch & Company.

Di Scipio, A. (2003). 'Sound is the interface ': from interactive to ecosystemic signal processing. *Organised Sound*, *8*(3), 269–277.

Djajadiningrat, T., Matthews, B., & Stienstra, M. (2007). Easy doesn't do it: skill and expression in tangible aesthetics. *Personal and Ubiquitous Computing*, *11*(8), 657–676. http://doi.org/10.1007/s00779-006-0137-9

Dobkin, B. H. (2004). Strategies for stroke rehabilitation. *The Lancet. Neurology*, *3*(9), 528–36. http://doi.org/10.1016/S1474-4422(04)00851-8

Dourish, P. (2001). *Where the action is: Foundations of embodied interaction*. MIT Press.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition*, *74*, 1–32.

Dreyfus, H. L. (1996). The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment Skill Acquisition: The Establishment of the Intentional Arc. *The Electronic Journal of Analytic Philosophy*, *4*, 1–13.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology Section A*, *58*(8), 1376–1389. http://doi.org/10.1080/02724980443000610

Dubus, G., & Bresin, R. (2013). A systematic review of mapping strategies for the sonification of physical quantities. *PloS One*, *8*(12), e82491. http://doi.org/10.1371/journal.pone.0082491

Dyer, J., Stapleton, P., & Rodger, M. W. M. (2015). Sonification as concurrent augmented feedback for motor skill learning and the importance of mapping design. *The Open Psychology Journal*, *8*(3), 1–11.

Dyer, J., Stapleton, P., & Rodger, M. W. M. (2017). Transposing musical skill: sonification of movement as concurrent augmented feedback enhances learning in a bimanual task. *Psychological Research*, *81*(4), 850–862. http://doi.org/10.1007/s00426-016-0775-0

Effenberg, A. O. (2005). Movement Sonification: Effects on Perception and Action. *IEEE Multimedia*, *12*(2), 53–59. http://doi.org/10.1109/MMUL.2005.31

Effenberg, A. O., Fehse, U., Schmitz, G., Krueger, B., & Mechling, H. (2016). Movement Sonification: Effects on Motor Learning beyond Rhythmic Adjustments. *Frontiers in Neuroscience*, *10*(May). http://doi.org/10.3389/fnins.2016.00219

Effenberg, A. O., Schmitz, G., Baumann, F., Rosenhahn, B., & Kroeger, D. (2015). SoundScript – Supporting the Acquisition of Character Writing by Multisensory Integration. *The Open Psychology Journal*, *8*(3), 230–237.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. http://doi.org/10.1037/0033-295X.100.3.363

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in*

Cognitive Sciences, *8*(4), 162–9. http://doi.org/10.1016/j.tics.2004.02.002

Fitch, W., & Kramer, G. (1994). Sonifying the Body Electric: Superiority of an Auditory over a Visual Display in a Complex, Multivariate System. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces* (pp. 307–325). Reading, MA: Addison-Wesley.

Fitts, P. M., & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole Publishing Company.

Flach, R., Knoblich, G., & Prinz, W. (2004). Recognizing one's own clapping: the role of temporal cues. *Psychological Research*, *69*(1–2), 147–156. http://doi.org/10.1007/s00426-003-0165-2

Flowers, J. H. (2005). Thirteen Years of Reflection on Auditory Graphing: Promises, Pitfalls, and Potential New Directions. In *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display, Limerick, Ireland, July 6-9*.

Flowers, J. H., & Grafel, D. C. (2002). Perception of Sonified Daily Weather Records. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, pp. 1579–1583). SAGE Publications. http://doi.org/10.1177/154193120204601711

Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A., & Pylyshyn, Z. (1981). How direct is visual perception?: Some reflections on Gibson's "Ecological Approach." *Cognition*, *9*, 139–196.

Fowler, C. A., & Turvey, M. T. (1978). Skill Acquisition: An event approach with special reference to searching for the optimum of a function of several variables. In G. E. Stelmach (Ed.), *Information Processing in Motor Control and Learning*. New York, New York, USA: New York: Academic Press.

Franz, E. A., & McCormick, R. (2010). Conceptual unifying constraints override sensorimotor interference during anticipatory control of bimanual actions. *Experimental Brain Research*, *205*(2), 273–82. http://doi.org/10.1007/s00221-010-2365-5

Franz, E. A., Zelaznik, H. N., Swinnen, S. P., & Walter, C. B. (2001). Spatial conceptual influences on the coordination of bimanual actions: when a dual task becomes a single task. *Journal of Motor Behavior*, *33*(1), 103–12. http://doi.org/10.1080/00222890109601906

Frid, E., Bresin, R., Alborno, P., & Elblaus, L. (2016). Interactive Sonification of Spontaneous Movement of Children—Cross-Modal Mapping and the Perception of

Body Movement Qualities through Sound. *Frontiers in Neuroscience*, *10*(November). http://doi.org/10.3389/fnins.2016.00521

Fyans, A. C., Gurevich, M., & Stapleton, P. (2010). Examining the Spectator Experience. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, (November), 451–454. http://doi.org/10.1080/10459880903217689

Gallagher, S. (2013). The socially extended mind. *Cognitive Systems Research*, *25–26*, 4–12. http://doi.org/10.1016/j.cogsys.2013.03.008

Gaver, W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, *5*(1), 1–29.

Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey - I. Somatotopy and the control of proximal movements. *Experimental Brain Research*, *71*(3), 475–490. http://doi.org/10.1007/BF00248741

Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York, NY: Appleton-Century-Crofts.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*, 1–41.

Gibson, E. J., & Walker, A. S. (1984). Development of Knowledge of Visual-Tactual Affordances of Substance. *Child Development*, *55*(2), 453–460.

Gibson, J. J. (1950). *The perception of the visual world.* Boston, MA: Houghton Mifflin.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.

Gibson, J. J. (1972). A Theory of Direct Visual Perception. In J. Royce & W. Rozenboom (Eds.), *The Psychology of Knowing*. New York, NY: Gordon & Breach.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. New York, NY: Taylor & Francis Group.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual Learning: Differentiation or Enrichment? *Psychological Review*, *62*(1), 32–41. http://doi.org/10.1037/h0048826

Giordano, B. L., & McAdams, S. (2006). Material identification of real impact sounds: effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, *119*(2), 1171–81.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. Y. (2011).

Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology*, *2*(AUG), 1–12. http://doi.org/10.3389/fpsyg.2011.00146

Gurevich, M., & Fyans, C. A. (2011). Digital Musical Interactions: Performer–system relationships and their perception by spectators. *Organised Sound*, *16*(2), 166–175. http://doi.org/10.1017/S1355771811000112

Gurevich, M., Stapleton, P., & Marquez-Borbon, A. (2010). Style and Constraint in Electronic Musical Instruments. In *NIME '10 Proceedings of the 2010 conference on New interfaces for musical expression* (pp. 106–111).

Heitger, M. H., Ronsse, R., Dhollander, T., Dupont, P., Caeyenberghs, K., & Swinnen, S. P. (2012). Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *NeuroImage*, *61*(3), 633–50. http://doi.org/10.1016/j.neuroimage.2012.03.067

Hendriks-Jansen, H. (1996). *Catching ourselves in the act: Situated activity, interactive emergence, evolution and human thought*. Cambridge, MA: MIT Press.

Henkelmann, C. (2007). *Improving the aesthetic quality of realtime motion data sonification, computer graphics, Technical Report CG-2007-4, University of Bonn*.

Hermann, T., Hunt, A., & Neuhoff, J. (Eds.). (2011). *The Sonification Handbook* (1st ed.). Berlin: Logos Publishing House.

Hill, H. (2015). Letter to the editor concerning the article "Effects of acoustic feedback training in elite-standard Para-Rowing" by Schaffert and Mattes (2015). *Journal of Sports Sciences*, *33*(15), 1637–1638. http://doi.org/10.1080/02640414.2014.1002103

Hirsh, I. J., & Watson, C. S. (1996). Auditory psychophysics and perception. *Annual Review of Psychology*, *47*, 461–84. http://doi.org/10.1146/annurev.psych.47.1.461

Holst, E., & Mittelstaedt, H. (1980). The principle of reafference: Interactions between the central nervous system and the peripheral organs. In P. C. Dodwell (Ed.), *Perceptual processing: Stimulus equivalence and pattern recognition* (pp. 41–72). New York, NY: Appleton-Century-Crofts.

Höner, O. (2011). Aiding Movement with Sonification in "Exercise, Play and Sport." In T. Hermann, A. Hunt, & J. Neuhoff (Eds.), *The Sonification Handbook* (1st ed.). Berlin: Logos Publishing House.

Hornecker, E. (2005). A Design Theme for Tangible Interaction: Embodied Facilitation. *ECSCW*.

Horsak, B., Iber, M., Bauer, K., Kiselka, A., Gorgas, A.-M., Dlapka, R., & Doppler, J. (2016). A wireless instrumented insole device for real-time sonification of gait. *Journal on Multimodal User Interfaces*. http://doi.org/10.1007/s12193-016-0216-9

Houben, M. M. J., Kohlrausch, A., & Hermes, D. J. (2004). Perception of the size and speed of rolling balls by sound. *Speech Communication*, *43*(4), 331–345. http://doi.org/10.1016/j.specom.2004.03.004

Hourcade, J. P., Mascher, S. L., Wu, D., & Pantoja, L. (2015). Look, My Baby Is Using an iPad! An Analysis of YouTube Videos of Infants and Toddlers Using Tablets. *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems*, *1*, 1915–1924. http://doi.org/10.1145/2702123.2702266

Hove, M. J., & Keller, P. E. (2010). Spatiotemporal Relations and Movement Trajectories in Visuomotor Synchronization. *Music Perception*, *28*(1), 15–26. http://doi.org/10.1525/mp.2010.28.1.15

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(5), 772–784. http://doi.org/10.1037/0096-1523.9.5.772

Humphrey, G. (1924). The Psychology of the Gestalt. *Journal of Educational Psychology*, *15*(7), 401–412.

Hunt, A., Wanderley, M. M., & Paradis, M. (2003). The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research*, *32*(4), 429–440. http://doi.org/10.1076/jnmr.32.4.429.18853

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., … Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*(6766), 192–5. http://doi.org/10.1038/35003194

Ingold, T. (1996). Situating Action V: The History and Evolution of Bodily Skills. *Ecological Psychology*, *8*(2), 171–182. http://doi.org/10.1207/s15326969eco0802

Ingold, T. (2000). *The Perception of the Environment*. Abingdon, UK: Taylor & Francis. http://doi.org/10.4324/9780203466025

Ingold, T. (2001). Beyond art and technology: the Anthropology of Skill. In H. Schiffer (Ed.), *Anthropological perspectives on technology* (pp. 17–33). Albuquerque: University of New Mexico Press.

Jacobs, D. M., & Michaels, C. F. (2007). Direct Learning. *Ecological Psychology*, *19*(4), 321–349. http://doi.org/10.1080/10407410701432337

Jensen, M. V., Buur, J., & Djajadiningrat, T. (2005). Designing the user actions in tangible interaction. In *Proceedings of the 4th decennial conference on Critical computing between sense and sensibility - CC '05* (pp. 9–18). New York, New York, USA: ACM Press. http://doi.org/10.1145/1094562.1094565

Jensenius, A. R., & Lyons, M. J. (Eds.). (2017). *A NIME Reader - Fifteen Years of New Interfaces* (1st ed.). Springer International Publishing. http://doi.org/10.1007/978-3-319-47214-0

Johnson, M. (2007). *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago, IL: University of Chigaco Press.

Johnson, M. (2015). Embodied understanding. *Frontiers in Psychology*, *6*(June), 875. http://doi.org/10.3389/fpsyg.2015.00875

Johnson, M., & Larson, S. (2003). "Something in the way she moves" - Metaphors of musical motion. *Metaphor and Symbol*, *18*(2), 63–84. http://doi.org/10.1207/S15327868MS1802_1

Kagerer, F. A., & Contreras-Vidal, J. L. (2009). Adaptation of sound localization induced by rotated visual feedback in reaching movements. *Experimental Brain Research*, *193*(2), 315–321. http://doi.org/10.1007/s00221-008-1630-3

Karplus, K., & Strong, A. (1983). Digital Synthesis of and Plucked-String Timbres. *Computer Music Journal*, *7*(2), 43–55.

Kaufer, S., & Chemero, A. (2015). *Phenomenology: An introduction*. Cambridge: Polity Press.

Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *The Americal Journal of Physiology*, *246*(6), 1000–1004.

Kelso, J. A. S., Scholz, J. P., & Schoner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters A*, *118*(6), 279–284.

Kennedy, D. M., Boyle, J. B., & Shea, C. H. (2013). The role of auditory and visual models in the production of bimanual tapping patterns. *Experimental Brain Research*, *224*(4), 507–18. http://doi.org/10.1007/s00221-012-3326-y

Kennedy, D. M., Wang, C., Panzer, S., & Shea, C. H. (2016). Continuous scanning trials:Transitioning through the attractor landscape. *Neuroscience Letters*, *610*, 66–72. http://doi.org/10.1016/j.neulet.2015.10.073

Kennedy, J. M. (1985). Convergence principle in blind people's pointing. *International Journal of Rehabilitation Research*, *8*(2), 207–210.

Kennel, C., Hohmann, T., & Raab, M. (2014). Action perception via auditory information: Agent identification and discrimination with complex movement sounds. *Journal of Cognitive Psychology*, *26*(2), 157–165. http://doi.org/10.1080/20445911.2013.869226

Kennel, C., Streese, L., Pizzera, A., Justen, C., Hohmann, T., & Raab, M. (2015). Auditory reafferences: the influence of real-time feedback on movement control. *Frontiers in Psychology*, *6*(January), 69. http://doi.org/10.3389/fpsyg.2015.00069

Kernodle, M. W., & Carlton, L. G. (1992). Information Feedback and the Learning of Multiple-Degree-of-Freedom Activities. *Journal of Motor Behavior*, *24*(2), 187–195. http://doi.org/10.1080/00222895.1992.9941614

Kiverstein, J., & Miller, M. (2015). The embodied brain: towards a radical embodied cognitive neuroscience. *Frontiers in Human Neuroscience*, *9*(May), 1–11. http://doi.org/10.3389/fnhum.2015.00237

Klapp, S. T., & Jagacinski, R. J. (2011). Gestalt principles in the control of motor action. *Psychological Bulletin*, *137*(3), 443–462. http://doi.org/10.1037/a0022361

Klapp, S. T., Nelson, J. M., & Jagacinski, R. J. (1998). Can people tap concurrent bimanual rhythms independently? *Journal of Motor Behavior*, *30*(4), 301–22. http://doi.org/10.1080/00222899809601346

Kleiman-Weiner, M., & Berger, J. (2006). The sound of one arm swinging: A model for multidimensional auditory display of physical motion. In *12th International Conference on Auditory Display (ICAD), London, UK, June 20-23*. International Community on Auditory Display.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing Sounds , Understanding Actions: Action Representation in Mirror Neurons. *Science*, *297*(5582), 846–848.

Kolarik, A. J., Cirstea, S., Pardhan, S., & Moore, B. C. J. (2014). A summary of research investigating echolocation abilities of blind and sighted humans. *Hearing Research*, *310*, 60–68. http://doi.org/10.1016/j.heares.2014.01.010

Konttinen, N., Mononen, K., Viitasalo, J. T., & Mets, T. (2004). The effects of augmented auditory feedback on psychomotor skill learning in precision shooting. *Journal of Sport & Exercise Psychology*, *26*, 306–316.

Koutsoklenis, A., & Papadopolous, K. (2011). Auditory Cues Used for Wayfinding in Urban

Environments by Individuals with Visual Impairments. *Journal of Visual Impairment & Blindness*, *105*(10), 703–14.

Kovacs, A. J., Buchanan, J. J., & Shea, C. H. (2009). Bimanual 1:1 with 90° continuous relative phase: difficult or easy! *Experimental Brain Research*, *193*, 129–136.

Kovacs, A. J., Buchanan, J. J., & Shea, C. H. (2010). Impossible is nothing: 5:3 and 4:3 multi-frequency bimanual coordination. *Experimental Brain Research*, *201*(2), 249–59. http://doi.org/10.1007/s00221-009-2031-y

Kovacs, A. J., & Shea, C. H. (2011). The learning of 90° continuous relative phase with and without Lissajous feedback: external and internally generated bimanual coordination. *Acta Psychologica*, *136*(3), 311–20. http://doi.org/10.1016/j.actpsy.2010.12.004

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N., & Neuhoff, J. (1999). *The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display*. Santa Fe, NM.

Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, *4*(JAN), 1–13. http://doi.org/10.3389/fpsyg.2013.01003

Lahav, A., Katz, T., Chess, R., & Saltzman, E. (2013). Improved motor sequence retention by motionless listening. *Psychological Research*, *77*(3), 310–9. http://doi.org/10.1007/s00426-012-0433-0

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience*, *27*(2), 308–14. http://doi.org/10.1523/JNEUROSCI.4822-06.2007

Lai, Q., & Shea, C. H. (1998). Generalized Motor Program (GMP) Learning: Effects of Reduced Frequency of Knowledge of Results and Practice Variability. *Journal of Motor Behavior*, *30*(1), 51–59. http://doi.org/10.1080/00222899809601322

Lakoff, G. (2012). Explaining embodied cognition results. *Topics in Cognitive Science*, *4*(4), 773–85. http://doi.org/10.1111/j.1756-8765.2012.01222.x

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York, NY: Basic Books.

Latash, M. L., Scholz, J. P., & Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sport Sciences Reviews*, *30*(1), 26–31.

Lebedev, M. a, & Wise, S. P. (2002). Insights into seeing and grasping: distinguishing the neural correlates of perception and action. *Behavioral and Cognitive Neuroscience*

Reviews, *1*(2), 108–129. http://doi.org/10.1177/15382302001002002

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception, 5*(4), 437–59.

Lelièvre-Desmas, M., Chollet, S., Abdi, H., & Valentin, D. (2015). Becoming a beer expert: Is simple exposure with feedback sufficient to learn beer categories? *Acta Psychologica*, *161*, 95–103. http://doi.org/10.1016/j.actpsy.2015.08.003

Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge, MA: MIT Press.

Lotze, M., Scheler, G., Tan, H.-R. ., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *NeuroImage*, *20*(3), 1817–1829. http://doi.org/10.1016/j.neuroimage.2003.07.018

Magill, R. A. (2011). *Motor Learning and Control: Concepts and Applications* (9th ed.). New York, NY: McGraw-Hill.

Magnusson, T. (2009). Of Epistemic Tools: musical instruments as cognitive extensions. *Organised Sound*, *14*(2), 168. http://doi.org/10.1017/S1355771809000272

Magnusson, T., & Hurtado de Mendieta, E. (2007). The acoustic, the digital and the body: A survey on musical instruments. *Proceedings of the 2007 International Conference on New Instruments for Musical Expression (NIME-2007)*, 94–99. http://doi.org/10.1145/1279740.1279757

Maslovat, D., Brunke, K. M., Chua, R., & Franks, I. M. (2009). Feedback effects on learning a novel bimanual coordination pattern: support for the guidance hypothesis. *Journal of Motor Behavior*, *41*(1), 45–54. http://doi.org/10.1080/00222895.2009.10125923

Maulucci, R. A., & Eckhouse, R. H. (2001). Retraining reaching in chronic stroke with real-time auditory feedback. *NeuroRehabilitation*, *16*(3), 171–82.

McCabe, K., & Rangwalla, A. (1994). Auditory Display of Computational Fluid Dynamics Data. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification and Auditory Interfaces* (pp. 321–340). Reading, MA: Addison-Wesley.

McGee, R., & Rogers, D. (2016). Musification of seismic data. In *Proceedings of the 22nd International Conference on Auditory Display (ICAD-2016).*

Mechsner, F., Kerzel, D., Knoblich, È., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature*, *414*, 69–73.

Menzer, F., Brooks, A., Halje, P., Faller, C., Vetterli, M., & Blanke, O. (2010). Feeling in control of your footsteps: Conscious gait monitoring and the auditory consequences of footsteps. *Cognitive Neuroscience*, *1*(3), 184–192. http://doi.org/10.1080/17588921003743581

Michaels, C. F., & Carello, C. (1981). *Direct Perception*. Englewood Cliffs, New Jersey: Prentice-Hall.

Misenheimer, R., & Landreth, C. (1993). Caustic Sky. In R. Bargar, B. Evans, & C. Scaletti (Eds.), *An Introduction to Data Sonification*. Anaheim, CA: ACM Press.

Mithen, S. (2005). *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld & Nichoson.

Moneta, G. B., & Csikszentmihalyi, M. (1996). The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality*, *64*(2), 275–310. http://doi.org/10.1111/j.1467-6494.1996.tb00512.x

Mononen, K., Viitasalo, J. T., Konttinen, N., & Era, P. (2003). The effects of augmented kinematic feedback on motor skill learning in rifle shooting. *Journal of Sports Sciences*, *21*(10), 867–76. http://doi.org/10.1080/0264041031000101944

Mossio, M., & Taraborelli, D. (2008). Action-dependent perceptual invariants: from ecological to sensorimotor approaches. *Consciousness and Cognition*, *17*(4), 1324–40. http://doi.org/10.1016/j.concog.2007.12.003

Murgia, M., & Galmonte, A. (2015). The Role of Sound in Motor Perception and Execution. *The Open Psychology Journal*, *8*(3), 171–173. http://doi.org/10.2174/1874350101508010171

Murgia, M., Hohmann, T., Galmonte, A., Raab, M., & Agostini, T. (2012). Recognising one's own motor actions through sound: The role of temporal factors. *Perception, 41*(8), 976–987. http://doi.org/10.1068/p7227

Nagel, T. (1974). What Is It Like to Be A Bat? *The Philosophical Review*, *83*(4), 435–450. http://doi.org/10.2307/2183914

Newell, K. M. (1974). Knowledge of results and motor learning. *Journal of Motor Behaviour*, *6*(4), 235–244.

Newell, K. M., McDonald, P. V., & Kugler, P. N. (1991). The perceptual-motor workspace and the acquisition of skill. In *Tutorials in motor neuroscience* (pp. 95–108). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-011-3626-6_9

Newell, K. M., & van Emmerik, R. E. A. (1989). The acquisition of coordination: Preliminary analysis of learning to write. *Human Movement Science*, *8*(1), 17–32.

http://doi.org/10.1016/0167-9457(89)90021-3

Noe, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

Norman, D. (2002). *The design of everyday things*. New York, NY: Basic Books.

Norman, S. J., Ryan, J., & Waisvisz, M. (1998). Touchstone. In *Touch*. STEIM.

Nudo, R. J. (2006). Mechanisms for recovery of motor function following cortical damage. *Current Opinion in Neurobiology*, *16*(6), 638–44. http://doi.org/10.1016/j.conb.2006.10.004

Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Øritsland, T., & Buur, J. (2000). Taking the best from a company history-designing with interaction styles. In *DIS '00 Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 27–38).

Oscari, F., Secoli, R., Avanzini, F., Rosati, G., & Reinkensmeyer, D. J. (2012). Substituting auditory for visual feedback to adapt to altered dynamic and kinematic environments during reaching. *Experimental Brain Research*, *221*(1), 33–41. http://doi.org/10.1007/s00221-012-3144-2

Overbeeke, K., Djajadiningrat, T., Hummels, C., & Wensveen, S. (2000). Beauty in usability: Forget about ease of use! In W. Green & P. Jordan (Eds.), *Pleasure with Products, Beyond Usability*. New York, NY: Taylor & Francis.

Ozgen, E., & Davies, I. R. L. (2002). Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology. General*, *131*(4), 477–493. http://doi.org/10.1037/0096-3445.131.4.477

Park, J. H., Shea, C. H., & Wright, D. L. (2000). Reduced-frequency concurrent and terminal feedback: a test of the guidance hypothesis. *Journal of Motor Behavior*, *32*(3), 287–96. http://doi.org/10.1080/00222890009601379

Park, S.-W., & Sternad, D. (2015). Robust retention of individual sensorimotor skill after self-guided practice. *Journal of Neurophysiology*, *113*(7), 2635–2645. http://doi.org/10.1152/jn.00884.2014

Pfordresher, P. Q. (2009). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, *2*(2), 183–198. http://doi.org/10.2478/v10053-008-0054-8

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533–546. http://doi.org/10.1016/j.cognition.2006.11.006

Pick, H. (1992). Eleanor J. Gibson: Learning to perceive and perceiving to learn. *Developmental Psychology*, *28*(3), 787–794.

Pizzera, A., & Hohmann, T. (2015). Acoustic Information During Motor Control and Action Perception : A Review. *The Open Psychology Journal*, *8*(3), 183–191. http://doi.org/10.2174/1874350101508010183

Ramachandran, S., & Hubbard, E. M. (2001). Synaesthesia — A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Rath, M., & Schleicher, R. (2008). On the Relevance of Auditory Feedback for Quality of Control in a Balancing Task. *Acta Acustica United with Acustica*, *94*(1), 12–20. http://doi.org/10.3813/AAA.918003

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York, NY: Oxford University Press.

Reeves, S., Benford, S., O'Malley, C., & Fraser, M. (2005). Designing the spectator experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05*, 741. http://doi.org/10.1145/1054972.1055074

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual Span in Expert Chess Players: Evidence From Eye Movements. *Psychological Science*, *12*(1), 48–55.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). TO SEE OR NOT TO SEE : The Need for Attention to Perceive Changes in Scenes. *Psychological Science*, *8*(5), 368–373. http://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Repp, B. H. (1993). Music as Motion: A Synopsis of Alexander Truslit's (1938) "Gestaltung und Bewegung in der Musik." *Psychology of Music*, *21*(1), 48–72. http://doi.org/10.1177/030573569302100104

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.

Repp, B. H., & Knoblich, G. (2004). How Pianists Recognize Their Own Performances. *Psychological Science*, *15*(9), 604–609. http://doi.org/10.1111/j.0956-7976.2004.00727.x

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of*

Experimental Psychology. Human Perception and Performance, *28*(5), 1085–1099.

Rietveld, E., & Kiverstein, J. (2014). A Rich Landscape of Affordances. *Ecological Psychology*, *26*(4), 325–352. http://doi.org/10.1080/10407413.2014.958035

Robertson, J. V. G., Hoellinger, T., Lindberg, P., Bensmail, D., Hanneton, S., & Roby-Brami, A. (2009). Effect of auditory feedback differs according to side of hemiparesis: a comparative pilot study. *Journal of Neuroengineering and Rehabilitation*, *6*(45). http://doi.org/10.1186/1743-0003-6-45

Robinson, L. E. (2011). The relationship between perceived physical competence and fundamental motor skills in preschool children. *Child: Care, Health and Development*, *37*(4), 589–596. http://doi.org/10.1111/j.1365-2214.2010.01187.x

Roddy, S., & Bridges, B. (2016). Sounding human with data: The role of embodied conceptual metaphors and aesthetics in representing and exploring data sets. In *Music Technology Workshop*. Dublin.

Roddy, S., & Furlong, D. (2014). Embodied Aesthetics in Auditory Display. *Organised Sound*, *19*(1), 70–77. http://doi.org/10.1017/S1355771813000423

Roddy, S., & Furlong, D. (2015a). Embodied Auditory Display Affordances. In *Sound and Music Computing*.

Roddy, S., & Furlong, D. (2015b). Sonification listening: An empirical embodied approach. In *Proceedings of the 21st International conference on Auditory Display, At Graz, Austria* (pp. 181–187).

Rodger, M. W. M. (2010). *Musicians' Body Movements in Musical Skill Acquisition*. Queen's University Belfast.

Rodger, M. W. M., & Craig, C. M. (2011). Timing movements to interval durations specified by discrete or continuous sounds. *Experimental Brain Research*, *214*(3), 393–402. http://doi.org/10.1007/s00221-011-2837-2

Rodger, M. W. M., & Craig, C. M. (2014). Moving with beats and loops: The structure of auditory events and sensorimotor timing. In M. Aramaki, O. Derrien, R. Kronland-Martinet, & S. Ystad (Eds.), *Sound, Music, and Motion* (Vol. 8905, pp. 204–217). Springer International Publishing. http://doi.org/10.1007/978-3-319-12976-1

Rodger, M. W. M., & Craig, C. M. (2016). Beyond the Metronome: Auditory Events and Music May Afford More than Just Interval Durations as Gait Cues in Parkinson's Disease. *Frontiers in Neuroscience*, *10*(June). http://doi.org/10.3389/fnins.2016.00272

Rodger, M. W. M., Craig, C. M., & O'Modhrain, S. (2012). Expertise is perceived from both sound and body movement in musical performance. *Human Movement Science*, *31*(5), 1137–1150. http://doi.org/10.1016/j.humov.2012.02.012

Rodger, M. W. M., Young, W. R., & Craig, C. M. (2014). Synthesis of walking sounds for alleviating gait disturbances in Parkinson’s disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(3), 543–8. http://doi.org/10.1109/TNSRE.2013.2285410

Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., & Swinnen, S. P. (2011). Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cerebral Cortex*, *21*(6), 1283–94. http://doi.org/10.1093/cercor/bhq209

Rosati, G., Oscari, F., Spagnol, S., Avanzini, F., & Masiero, S. (2012). Effect of task-related continuous auditory feedback during learning of tracking motion exercises. *Journal of Neuroengineering and Rehabilitation*, *9*(79), 1–13. http://doi.org/10.1186/1743-0003-9-79

Rosati, G., Rodà, A., Avanzini, F., & Masiero, S. (2013). On the role of auditory feedback in robot-assisted movement training after stroke: review of the literature. *Computational Intelligence and Neuroscience*, *2013*(586138), 1–15. http://doi.org/10.1155/2013/586138

Rosenblum, L. D. (2008). Speech perception as a multimodal phenomenon. *Current Directions in Psychological Science*, *17*(6), 405–409. http://doi.org/10.1111/j.1467-8721.2008.00615.x

Rosenblum, L. D., Dias, J. W., & Dorsi, J. (2016). The supramodal brain: implications for auditory perception. *Journal of Cognitive Psychology*, *5911*(June), 1–23. http://doi.org/10.1080/20445911.2016.1181691

Rosenblum, L. D., Miller, R. M., & Sanchez, K. (2007). Lip-read me now, hear me better later: Cross-modal transfer of talker-familiarity effects: Research report. *Psychological Science*, *18*(5), 392–396. http://doi.org/10.1111/j.1467-9280.2007.01911.x

Runeson, S. (1977). On the possibility of “smart” perceptual mechanisms. *Scandinavian Journal of Psychology*, *18*(3), 172–179. http://doi.org/10.1111/j.1467-9450.1977.tb00274.x

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113–29. http://doi.org/10.1016/j.cognition.2005.01.004

Salgado-Montejo, A., Marmolejo-Ramos, F., Alvarado, J. A., Arboleda, J. C., Suarez, D. R., & Spence, C. (2016). Drawing sounds: representing tones and chords spatially. *Experimental Brain Research*, *234*(12), 3509–3522. http://doi.org/10.1007/s00221-016-4747-9

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of Results and Motor Learning : A Review and Critical Reappraisal. *Psychological Bulletin*, *95*(3), 355–386.

Sanchez, K., Dias, J. W., & Rosenblum, L. D. (2013). Experience with a Talker Can Transfer Across Modalities to Facilitate Lipreading. *Attention, Perception, & Psychophysics*, *75*(7), 1359–1365. http://doi.org/10.3758/s13414- 013-0534-x

Saue, S. (2000). A model for interaction in exploratory sonification displays. In *Proceedings of the 6th International Conference on Auditory Display (ICAD), Atlanta, GA, April 2-5*. International Community for Auditory Display.

Scaletti, C., & Craig, A. B. (1991). Using sound to extract meaning from complex data. In *Proceedings of SPIE; Extracting meaning from complex data: processing, display, interaction II* (pp. 207–219). http://doi.org/10.1117/12.44397

Schaffert, N., & Gehret, R. (2013). Testing different versions of functional sonification as acoustic feedback for rowing. In *International Conference on Auditory Display (ICAD), Poland* (pp. 331–335).

Schaffert, N., & Mattes, K. (2015). Effects of acoustic feedback training in elite-standard Para-Rowing. *Journal of Sports Sciences*, *33*(4), 411–8. http://doi.org/10.1080/02640414.2014.946438

Schaffert, N., Mattes, K., Barrass, S., & Effenberg, A. O. (2009). Exploring function and aesthetics in sonifications for elite sports. In *Proceedings of The Second International Conference on Music Communication Science* (pp. 3–4). Sydney, Australia.

Schellenberg, E. G., & Hallam, S. (2005). Music listening and cognitive abilities in 10- and 11-year-olds: the blur effect. *Annals of the New York Academy of Sciences*, *1060*, 202–9. http://doi.org/10.1196/annals.1360.013

Schiavio, A., & Altenmüller, E. O. (2015). Exploring music-based rehabilitation for Parkinsonism through embodied cognitive science. *Frontiers in Neurology*, *6*(OCT), 1–15. http://doi.org/10.3389/fneur.2015.00217

Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: effects of modality, trajectory, and gender. *Journal of Experimental Psychology : Human Perception and Performance*, *16*(2), 303–316. http://doi.org/10.1037/0096-1523.16.2.303

Schmidt, R. A. (1991). Frequent Augmented Feedback Can Degrade Learning: Evidence and Interpretations. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in Motor Neuroscience* (pp. 59–75). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-011-3626-6

Schmidt, R. A., & Wulf, G. (1997). Continuous Concurrent Feedback Degrades Skill Learning: Implications for Training and Simulation. *Human Factors*, *39*(4), 509–525.

Schmidt, R. A., Young, D. E., Swinnen, S., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: support for the guidance hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*(2), 352–9.

Schmitz, G., & Bock, O. (2014). A comparison of sensorimotor adaptation in the visual and in the auditory modality. *PLoS ONE*, *9*(9). http://doi.org/10.1371/journal.pone.0107834

Schmitz, G., Kroeger, D., & Effenberg, A. O. (2014). A mobile sonification system for stroke rehabilitation. In *The 20th International Conference on Auditory Display (ICAD-2014)*. New York, NY.

Schmitz, G., Mohammadi, B., Hammer, A., Heldmann, M., Samii, A., Münte, T. F., & Effenberg, A. O. (2013). Observation of sonified movements engages a basal ganglia frontocortical network. *BMC Neuroscience*, *14*(1), 32. http://doi.org/10.1186/1471-2202-14-32

Scholz, D. S., Rhode, S., Großbach, M., Rollnik, J., & Altenmüller, E. O. (2015). Moving with music for stroke rehabilitation: a sonification feasibility study. *Annals of the New York Academy of Sciences*, *1337*, 69–76. http://doi.org/10.1111/nyas.12691

Scholz, D. S., Wu, L., Pirzer, J., Schneider, J., Rollnik, J. D., Großbach, M., & Altenmüller, E. O. (2014). Sonification as a possible stroke rehabilitation strategy. *Frontiers in Neuroscience*, *8*(October), 332. http://doi.org/10.3389/fnins.2014.00332

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424. http://doi.org/10.1016/B978-1-4832-1446-7.50007-8

Sedda, A., Monaco, S., Bottini, G., & Goodale, M. A. (2011). Integration of visual and auditory information for hand actions: Preliminary evidence for the contribution of natural sounds to grasping. *Experimental Brain Research*, *209*(3), 365–374. http://doi.org/10.1007/s00221-011-2559-5

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, *77*(5), 406–418. http://doi.org/10.1037/h0029790

Shea, C. H., Buchanan, J. J., & Kennedy, D. M. (2016). Perception and action influences on discrete and reciprocal bimanual coordination. *Psychonomic Bulletin & Review*, *23*(2), 361–386. http://doi.org/10.3758/s13423-015-0915-3

Shea, C. H., Wulf, G., Park, J. H., & Gaunt, B. (2001). Effects of an auditory model on the learning of relative and absolute timing. *Journal of Motor Behavior*, *33*(May 2015), 127–138. http://doi.org/10.1080/00222890109603145

Shinar, D., Meir, M., & Ben-Shoham, I. (1998). How Automatic Is Manual Gear Shifting? *Human Factors*, *40*(4), 647–654. http://doi.org/10.1518/001872098779649346

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(1), 70–5. http://doi.org/10.1073/pnas.1209023110

Sigrist, R., Fox, S., Riener, R., & Wolf, P. (2016). Benefits of Crank Moment Sonification in Cycling. *Procedia Engineering*, *147*, 513–518. http://doi.org/10.1016/j.proeng.2016.06.230

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013a). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic Bulletin & Review*, *20*(1), 21–53. http://doi.org/10.3758/s13423-012-0333-8

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013b). Terminal feedback outperforms concurrent visual, auditory, and haptic feedback in learning a complex rowing-type task. *Journal of Motor Behavior*, *45*(6), 455–72. http://doi.org/10.1080/00222895.2013.826169

Sigrist, R., Schellenberg, J., Rauter, G., Broggi, S., Riener, R., & Wolf, P. (2011). Visual and Auditory Augmented Concurrent Feedback in a Complex Motor Task. *Presence: Teleoperators and Virtual Environments*, *20*(1), 15–32. http://doi.org/10.1162/pres_a_00032

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267. http://doi.org/10.1016/S1364-6613(97)01080-2

Sloboda, J. (1986). Computers and Cognition. In A. Gellatly (Ed.), *The Skilful Mind: An introduction to Cognitive Psychology*. Milton Keynes: Open University Press.

Smith, D. R., & Walker, B. N. (2002). Tick-marks, Axes, and Labels: The Effects of Adding Context to Auditory Graphs. In *International Conference on Auditory Display* (pp. 1–6).

Snapp-Childs, W., Wilson, A. D., & Bingham, G. P. (2015). Transfer of learning between unimanual and bimanual rhythmic movement coordination: transfer is a function of the task dynamic. *Experimental Brain Research*, *233*(7), 2225–38. http://doi.org/10.1007/s00221-015-4292-y

Soderstrom, N. C., & Bjork, R. a. (2015). Learning versus performance: an integrative review. *Perspectives on Psychological Science*, *10*(2), 176–199. http://doi.org/10.1177/1745691615569000

Speeth, S. D. (1961). Seismometer Sounds. *The Journal of the Acoustical Society of America*, *33*(7), 909. http://doi.org/10.1121/1.1908843

St Pierre, M., & Droumeva, M. (2016). Sonifying for public engagement: A context-based model for sonifying air pollution data. In *Proceedings of the 22nd International Conference on Auditory Display (ICAD-2016)*. Canberra, Australia.

Stapleton, P., & Davis, T. (2012). Ambiguous Devices. Retrieved from http://pure.qub.ac.uk/portal/en/publications/ambiguous-devices(275b8cb1-8c5e-4790-a233-48e5b1210503).html

Steenson, C., & Rodger, M. W. M. (2015). Bringing sounds into use : Thinking of sounds as materials and a sketch of auditory affordances. *The Open Psychology Journal*, *8*(3).

Stienstra, J., Overbeeke, K., & Wensveen, S. (2011). Embodying complexity through movement sonification. In *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity* (pp. 39–44). New York, New York, USA: ACM Press. http://doi.org/10.1145/2037296.2037310

Stoffregen, T. A. (2003). Affordances as Properties of the Animal-Environment System. *Ecological Psychology*, *15*(2), 115–134. http://doi.org/10.1207/S15326969ECO1502_2

Stoffregen, T. A., & Bardy, B. G. (2001). On specification and the senses. *The Behavioral and Brain Sciences*, *24*(2), 195–213. http://doi.org/10.1017/S0140525X01003946

Straebel, V. (2010). The Sonification Metaphor in Instrumental Music and Sonifications's Romantic Implications. In *Proceedings of ICAD 10-Sixteenth Meeting of the International Conference on Auditory Display* (pp. 287–294).

Sturm, B. L. (2000). Sonification of particle systems via de Broglie's hypothesis. In *Proceedings of the 6th International Conference on Auditory Display (ICAD 2000)*.

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, *76*(3), 373–82. http://doi.org/10.1007/s00426-011-0346-3

Suchman, L. A. (1987). *Plans and Situated Actions: The Problem of Human Machine Communication*. New York, NY: Cambridge University Press.

Summers, J. J., Rosenbaum, D. A., Burns, B. D., & Ford, S. K. (1993). Production of polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 416–428.

Supper, A. (2014). Sublime frequencies: The construction of sublime listening experiences in the sonification of scientific data. *Social Studies of Science*, *44*(1), 34–58. http://doi.org/10.1177/0306312713496875

Swinnen, S. P., & Gooijers, J. (2015). Bimanual Coordination. In A. W. Toga (Ed.), *Brain Mapping: An Encyclopedic Reference* (pp. 475–482). Academic Press: Elsevier. http://doi.org/10.1016/B978-0-12-397025-1.00030-0

Tajadura-jiménez, A., Basia, M., Deroy, O., Fairhurst, M., Marquardt, N., & Bianchi-Berthouze, N. (2015). As Light as your Footsteps: Altering Walking Sounds to Change Perceived Body Weight, Emotional State and Gait. In *Wellness & Wearables, Proceedings of CHI 2015, Seoul, Korea* (pp. 2943–2952).

Taylor, J. E. T., & Witt, J. K. (2015). Listening to music primes space: pianists, but not novices, simulate heard actions. *Psychological Research*, *79*(2), 175–82. http://doi.org/10.1007/s00426-014-0544-x

Thomas, K. T., & Thomas, J. R. (1994). Developing expertise in sport: The relation of knowledge and performance. *International Journal of Sport Psychology*, *25*(3), 295–312.

Tinazzi, M., Fiaschi, A., Frasson, E., Fiorio, M., Cortese, F., & Aglioti, S. M. (2002). Deficits of temporal discrimination in dystonia are independent from the spatial distance between the loci of tactile stimulation. *Movement Disorders : Official Journal of the Movement Disorder Society*, *17*(2), 333–8. http://doi.org/10.1002/mds.10019

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*(11), 1226–35. http://doi.org/10.1038/nn963

Todorov, E., Shadmehr, R., & Bizzi, E. (1997). Augmented Feedback Presented in a Virtual Environment Accelerates Learning of a Difficult Motor Task. *Journal of Motor Behavior*, *29*(2), 147–158. http://doi.org/10.1080/00222899709600829

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981). Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn. *Cognition*, *9*, 237–304. http://doi.org/10.1016/0010-0277(81)90002-0 van Dijk, L., Withagen, R., & Bongers, R. M. (2015). Information without content: a Gibsonian reply to enactivists' worries. *Cognition*, *134*, 210–4. http://doi.org/10.1016/j.cognition.2014.10.012

van Dinther, R., & Patterson, R. D. (2006). Perception of acoustic scale and size in musical instrument sounds. *The Journal of the Acoustical Society of America*, *120*(4), 2158–76.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception*, *30*(4), 349–359. http://doi.org/10.1525/mp.2013.30.4.34

van Vugt, F. T., & Tillmann, B. (2015). Auditory feedback in error-based learning of motor regularity. *Brain Research*, *1606*, 54–67. http://doi.org/10.1016/j.brainres.2015.02.026

Vander Linden, D. W., Cauraugh, J. H., & Greene, T. A. (1993). The effect of frequency of kinetic feedback on learning an isometric force production task in nondisabled subjects. *Physical Therapy*, *73*(2), 79–87.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vereijken, B., Emmerik, R. E. A. van, Whiting, H. T. A., & Newell, K. M. (1992). Free(z)ing Degrees of Freedom in Skill Acquisition. *Journal of Motor Behavior*, *24*(1), 133–142. http://doi.org/10.1080/00222895.1992.9941608

Verrel, J., Pologe, S., Manselle, W., Lindenberger, U., & Woollacott, M. (2013). Coordination of degrees of freedom and stabilization of task variables in a complex motor skill: expertise-related differences in cello bowing. *Experimental Brain Research*, *224*(3), 323–34. http://doi.org/10.1007/s00221-012-3314-2

Verwey, W. B., Shea, C. H., & Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychon Bull Rev*, *22*, 54–77. http://doi.org/10.3758/s13423-014-0773-4

Vickers, P. (2012). Ways of listening and modes of being: Electroacoustic auditory display. *Journal of Sonic Studies*, *2*(1).

Vinken, P. M., Kröger, D., Fehse, U., Schmitz, G., Brock, H., & Effenberg, A. O. (2013). Auditory Coding of Human Movement Kinematics. *Multisensory Research*, *26*(6), 533–552. http://doi.org/10.1163/22134808-00002435

Walker, B. N., & Kramer, G. (2004). Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping, and Meaning Making. In G. Kramer (Ed.), *Ecological Psychoacoustics* (pp. 150–175). San Diego, CA: Elsevier Academic.

Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays. *ACM Transactions on Applied Perception*, *2*(4), 407–412. http://doi.org/10.1145/1101530.1101534

Walker, E., & Nowacki, A. S. (2011). Understanding equivalence and noninferiority testing. *Journal of General Internal Medicine*, *26*(2), 192–196. http://doi.org/10.1007/s11606-010-1513-8

Wallace, S. A., & Hagler, R. W. (1979). Knowledge of Performance and the Learning of a Closed Motor Skill. *Research Quarterly. American Alliance for Health, Physical Education, Recreation and Dance*, *50*(2), 265–271. http://doi.org/10.1080/10671315.1979.10615609

Wallis, R. I., Ingalls, T., Rikakis, T., Olson, L., Chen, Y., Xu, W., & Sundaram, H. (2007). Real-Time Sonification of Movement for an Immersive Stroke Rehabilitation. In *Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada, June 26-29*.

Wang, C., Kennedy, D. M., Boyle, J. B., & Shea, C. H. (2013). A guide to performing difficult bimanual coordination tasks: just follow the yellow brick road. *Experimental Brain Research*, *230*(1), 31–40. http://doi.org/10.1007/s00221-013-3628-8

Ward, N. (2013). *Effortful Interaction : A New Paradigm for the Design of Digital Musical Instruments*. Queen’s University Belfast.

Warren, W. H. (1984). Perceiving affordances: visual guidance of stair climbing. *Journal of Experimental Psychology. Human Perception and Performance*, *10*(5), 683–703.

Warren, W. H. (1998). Visually Controlled Locomotion: 40 years Later. *Ecological Psychology*, *10*(3), 177–219. http://doi.org/10.1207/s15326969eco103&4_3

Watanabe, K., & Shimojo, S. (2001). When Sound Affects Vision: Effects of Auditory Grouping on Visual Motion Perception. *Psychological Science*, *12*(2), 109–116. http://doi.org/10.1111/1467-9280.00319

Waters, S. (2007). Performance Ecosystems: Ecological approaches to musical interaction. In *Proceedings of the Electroacoustic Music Studies Network Conference*.

Weiskopf, D. A. (2011). Models and mechanisms in psychological explanation. *Synthese*, *183*(3), 313–338. http://doi.org/10.1007/s11229-011-9958-9

Wessel, D., & Wright, M. (2002). Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal*, *26*(3), 11–22. http://doi.org/10.1162/014892602320582945

Willcock, I. (2006). Composing without Composers? : Creation, control and individuality in computer-based algorithmic composition. In C. Mahnkopf, F. Cox, & W. Schurig (Eds.), *Electronics in New Music: New Music and Aesthetics in the 21st Century* (pp. 221–235). Hofheim: Wolke Verlag.

Wilson, A. D., Collins, D. R., & Bingham, G. P. (2005). Perceptual coupling in rhythmic movement coordination: stable perception leads to stable action. *Experimental Brain Research*, *164*(4), 517–28. http://doi.org/10.1007/s00221-005-2272-3

Wilson, A. D., & Golonka, S. (2013). Embodied Cognition is Not What you Think it is. *Frontiers in Psychology*, *4*(February), 58. http://doi.org/10.3389/fpsyg.2013.00058

Wilson, A. D., & Golonka, S. (2015). Connecting the conceptual dots in embodied cognition: A commentary on "How body balance influences political party evaluations: a Wii balance board study." *Frontiers in Psychology*, *6*(June), 1–2. http://doi.org/10.3389/fpsyg.2015.00853

Wilson, A. D., Snapp-Childs, W., & Bingham, G. P. (2010). Perceptual learning immediately yields new stable motor coordination. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(6), 1508–14. http://doi.org/10.1037/a0020412

Wilson, A. D., Snapp-Childs, W., Coats, R., & Bingham, G. P. (2010). Learning a coordinated rhythmic movement with task-appropriate coordination feedback. *Experimental Brain Research*, *205*(4), 513–20. http://doi.org/10.1007/s00221-010-2388-y

Windsor, W. L., & de Bezenac, C. (2012). Music and affordances. *Musicae Scientiae*, *16*(1), 102–120. http://doi.org/10.1177/1029864911435734

Winstein, C. J., & Schmidt, R. A. (1991). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 677–691.

Withagen, R., & van Dijk, L. (2016). Temporalizing agency : Moving beyond on- and offline cognition. *Theory & Psychology*, *26*(1). http://doi.org/10.1177/0959354315596080

Witt, J. K., Linkenauger, S. A., & Wickens, C. (2016). Action-specific effects in perception and their potential applications. *Journal of Applied Research in Memory and Cognition*, *5*(1), 69–76. http://doi.org/10.1016/j.jarmac.2015.07.008

Wittwer, J. E., Webster, K. E., & Hill, K. (2013). Music and metronome cues produce different effects on gait spatiotemporal measures but not gait variability in healthy older adults. *Gait & Posture*, *37*(2), 219–22. http://doi.org/10.1016/j.gaitpost.2012.07.006

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience*, *12*(12), 739–51. http://doi.org/10.1038/nrn3112

Worrall, D. (2010). Parameter mapping sonic articulation and the perceiveing body. In *Proceedings of the 16th International Conference on Auditory Display* (pp. 207–214).

Wulf, G. (2013). Attentional focus and motor learning : a review of 15 years. *International Review of Sport and Exercise Psychology*, *6*(1), 77–104.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, *9*(2), 185–211. http://doi.org/10.3758/BF03196276

Wulf, G., Shea, C. H., & Lewthwaite, R. (2010). Motor skill learning and performance: a review of influential factors. *Medical Education*, *44*(1), 75–84. http://doi.org/10.1111/j.1365-2923.2009.03421.x

Young, W. R., Rodger, M. W. M., & Craig, C. M. (2013). Perceiving and reenacting spatiotemporal characteristics of walking sounds. *Journal of Experimental Psychology. Human Perception and Performance, 39*(2), 464–476. http://doi.org/10.1037/a0029402

Appendix A: Video recordings of sonification task performance

Bimanual object-manipulation task:

https://vimeo.com/219911401

Bimanual shape-tracing task with feedback used in experiment 2:

https://link.springer.com/article/10.1007/s00426-016-0775-0#SupplementaryMaterial

Modelled string synthesis sonification used in ‘Melodic’ conditions in experiments 3 & 4:

https://vimeo.com/219944175